GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:22:32 ; Search time 16 Seconds
(without alignments)
48.108 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 20
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	14	2	PH0804	T-cell receptor al
2	20	100.0	15	2	PH0797	T-cell receptor al
3	20	100.0	20	2	B44913	periplasmic flagel
4	20	100.0	22	2	S47206	T-cell receptor J-
5	20	100.0	22	2	F84018	hypothetical prote
6	20	100.0	23	2	S47192	T-cell receptor J-
7	20	100.0	24	2	C47689	flagellar core pro
8	20	100.0	25	2	B47689	flagellar core pro
9	20	100.0	26	2	S14035	hypothetical prote
10	20	100.0	26	2	S14036	hypothetical prote
11	20	100.0	26	2	S14037	hypothetical prote
12	20	100.0	26	2	S13989	hypothetical prote
13	20	100.0	28	2	S41774	ubiquinol-cytochro
14	20	100.0	28	2	PT0366	T-cell receptor be
15	20	100.0	30	2	A44913	34K core flagella
16	20	100.0	30	2	S74112	proline-rich antib
17	20	100.0	31	2	S53153	gene X protein - h
18	20	100.0	31	2	S53192	gene X protein - h
19	20	100.0	32	2	S39785	kallidin-releasing
20	20	100.0	32	2	B82421	hypothetical prote
21	20	100.0	34	2	F81044	hypothetical prote
22	20	100.0	34	2	H83722	hypothetical prote
23	20	100.0	36	2	A81164	hypothetical prote
24	20	100.0	36	2	A69326	hypothetical prote
25	20	100.0	38	2	A59185	photosystem II pro
26	20	100.0	38	2	D84227	hypothetical prote
27	20	100.0	38	2	H81579	hypothetical prote
28	20	100.0	39	2	A81151	hypothetical prote
29	20	100.0	39	2	B85990	hypothetical prote
30	20	100.0	40	2	S07969	T-cell receptor al
31	20	100.0	41	2	B27579	T-cell receptor be
32	20	100.0	41	2	D82691	hypothetical prote
33	20	100.0	41	2	E82544	hypothetical prote
34	20	100.0	42	2	G56271	lpfA 5'-region hyp
35	20	100.0	42	2	T07581	hypothetical prote
36	20	100.0	42	2	D81730	hypothetical prote
37	20	100.0	42	2	B82629	hypothetical prote
38	20	100.0	43	2	F81505	hypothetical prote
39	20	100.0	44	2	PL0091	Ig heavy chain V r
40	20	100.0	45	2	H83936	hypothetical prote
41	20	100.0	46	2	F95023	hypothetical prote
42	20	100.0	46	2	A99802	hypothetical prote
43	20	100.0	46	2	D84334	hypothetical prote
44	20	100.0	46	2	AG3569	hypothetical prote
45	20	100.0	47	2	S39358	cyclin kinase inhi
46	20	100.0	48	2	T20751	hypothetical prote
47	20	100.0	48	2	H84063	hypothetical prote
48	20	100.0	48	2	B28118	ubiquinol-cytochro
49	20	100.0	49	2	D81567	hypothetical prote
50	20	100.0	50	2	H90537	hypothetical prote
51	20	100.0	50	2	H82423	hypothetical prote
52	20	100.0	50	2	A69055	hypothetical prote
53	20	100.0	50	2	G82652	hypothetical prote
54	20	100.0	51	2	S64712	formin binding pro
55	20	100.0	52	2	AC3039	hypothetical prote
56	20	100.0	52	2	AB1904	hypothetical prote
57	20	100.0	53	2	G90776	hypothetical prote
58	20	100.0	54	2	T07247	hypothetical prote
59	20	100.0	54	2	A97904	hypothetical prote
60	20	100.0	55	2	T10344	hypothetical prote
61	20	100.0	55	2	H87442	hypothetical prote
62	20	100.0	55	2	AE0303	hypothetical prote
63	20	100.0	56	2	A29235	hemolymph trypsin
64	20	100.0	56	2	AD0739	probable bacteriop
65	20	100.0	56	2	T09177	ydaE protein - Esc
66	20	100.0	56	2	D82138	hypothetical prote
67	20	100.0	56	2	T07482	hypothetical prote
68	20	100.0	56	2	AB2413	hypothetical prote
69	20	100.0	57	2	AG3105	conserved hypothet
70	20	100.0	57	2	S16587	hypothetical prote
71	20	100.0	57	2	C87565	hypothetical prote
72	20	100.0	57	2	B95384	protein [imported
73	20	100.0	58	2	I53690	cytochrome P450 2B
74	20	100.0	58	2	H91110	hypothetical prote
75	20	100.0	58	2	AC0327	hypothetical prote
76	20	100.0	59	2	T07976	dihydrokaempferol
77	20	100.0	59	2	A36589	bactenecin 7 - bov
78	20	100.0	59	2	H86647	hypothetical prote
79	20	100.0	59	2	T30378	hypothetical prote
80	20	100.0	59	2	AC2752	hypothetical prote
81	20	100.0	60	2	T07272	hypothetical prote
82	20	100.0	60	2	D90971	hypothetical prote
83	20	100.0	60	2	D83610	hypothetical prote
84	20	100.0	60	2	D85744	unknown protein en
85	20	100.0	60	2	AF2783	hypothetical prote
86	20	100.0	61	1	N1NJ4E	short neurotoxin 4
87	20	100.0	61	2	D32021	bactericidin B-5P
88	20	100.0	61	2	S78741	protein YPR170w-a
89	20	100.0	61	2	A84004	hypothetical prote
90	20	100.0	61	2	G84029	hypothetical prote
91	20	100.0	62	2	S22564	finger protein - h
92	20	100.0	62	2	T08470	tryptophan-specifi
93	20	100.0	62	2	T03340	gene e27 protein -
94	20	100.0	62	2	E82630	hypothetical prote
95	20	100.0	63	1	CKMTA	cecropin A precurs
96	20	100.0	63	2	A29286	sucrose alpha-gluc
97	20	100.0	63	2	H84265	hypothetical prote
98	20	100.0	63	2	A71866	hypothetical prote
99	20	100.0	63	2	T09533	COX17 protein - hu
100	20	100.0	63	2	G72654	hypothetical prote

ALIGNMENTS

RESULT 1
PH0804
T-cell receptor alpha chain (L4) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0804
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0804
A;Molecule type: mRNA
A;Residues: 1-14 <CAS>
A;Cross-references: UNIPARC:UPI000017C77A; EMBL:X60913
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 100.0%; Score 20; DB 2; Length 14;
Best Local Similarity 37.5%; Pred. No. 2.5e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 6 QGGRALIF 13

RESULT 2
PH0797
T-cell receptor alpha chain (PF2.10.1 V-alpha-3.AR5) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PH0797
R;Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A;Title: T cell receptor genes in a series of class I major histocompatibility complex-r
allelic exclusion and antigen-specific repertoire.
A;Reference number: PH0746; MUID:92078846; PMID:1836010
A;Accession: PH0797
A;Molecule type: mRNA
A;Residues: 1-15 <CAS>
A;Cross-references: UNIPARC:UPI000017C780; EMBL:X60903
A;Experimental source: T lymphocyte
C;Keywords: T-cell receptor

Query Match 100.0%; Score 20; DB 2; Length 15;
Best Local Similarity 37.5%; Pred. No. 2.6e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 7 GNTRKLIF 14

RESULT 3
B44913
periplasmic flagellar core protein, 35.5K - Leptospira interrogans (fragment)
C;Species: Leptospira interrogans
C;Date: 01-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B44913; B41210
R;Trueba, G.A.; Bolin, C.A.; Zuerner, R.L.
J. Bacteriol. 174, 4761-4768, 1992
A;Title: Characterization of the periplasmic flagellum proteins of Leptospira interrogan
A;Reference number: A44913; MUID:92325069; PMID:1624463
A;Accession: B44913
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-20 <TRU>
A;Cross-references: UNIPROT:Q9R5K2; UNIPARC:UPI00000B3E62
A;Experimental source: sv. pomona type kennewicki
A;Note: sequence extracted from NCBI backbone (NCBIP:108222)

Query Match 100.0%; Score 20; DB 2; Length 20;
Best Local Similarity 37.5%; Pred. No. 3.5e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 11 FAHRTLKF 18

RESULT 4
S47206
T-cell receptor J-alpha wnVII.1 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C;Accession: S47206
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993
A;Reference number: S40133
A;Accession: S47206
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-22 <PLA>
A;Cross-references: UNIPARC:UPI0000116127; EMBL:X71036; NID:g507043; PIDN:CAA50353.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 100.0%; Score 20; DB 2; Length 22;
Best Local Similarity 37.5%; Pred. No. 3.8e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 5 TGRRALTF 12

RESULT 5
F84018
hypothetical protein BH2950 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: F84018
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: F84018
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-22 <STO>
A;Cross-references: UNIPROT:Q9K8Q3; UNIPARC:UPI00000C406A; GB:AP001517; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2950

Query Match 100.0%; Score 20; DB 2; Length 22;
Best Local Similarity 37.5%; Pred. No. 3.8e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 MKHRVLLF 8

RESULT 6
S47192
T-cell receptor J-alpha wnVII.2 - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Jul-1999
C;Accession: S47192
R;Plaza, A.; Kono, D.H.; Theofilopoulos, A.N.
submitted to the EMBL Data Library, February 1993

```
A;Reference number: S40133
A;Accession: S47192
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-23 <PLA>
A;Cross-references: UNIPARC:UPI0000116136; EMBL:X71051; NID:g506974; PIDN:CAA50368.1; PI
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

  Query Match             100.0%;  Score 20;  DB 2;  Length 23;
  Best Local Similarity   37.5%;  Pred. No. 4e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          6 TGRRALTF 13


RESULT 7
C47689
flagellar core protein, 34K - Treponema hyodysenteriae (fragment)
C;Species: Treponema hyodysenteriae
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: C47689
R;Koopman, M.B.; Baats, E.; van Vorstenbosch, C.J.; van der Zeijst, B.A.; Kusters, J.G.
J. Gen. Microbiol. 138, 2697-2706, 1992
A;Title: The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are composed c
A;Reference number: A47689; MUID:93139764; PMID:1487733
A;Contents: C5, Treponema
A;Accession: C47689
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-24 <KOO>
A;Cross-references: UNIPROT:Q7M132; UNIPARC:UPI000017A907
A;Note: sequence extracted from NCBI backbone (NCBIP:123402)

  Query Match             100.0%;  Score 20;  DB 2;  Length 24;
  Best Local Similarity   37.5%;  Pred. No. 4.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NAQRTLKF 18


RESULT 8
B47689
flagellar core protein, 37K - Treponema hyodysenteriae (fragment)
C;Species: Treponema hyodysenteriae
C;Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C;Accession: B47689
R;Koopman, M.B.; Baats, E.; van Vorstenbosch, C.J.; van der Zeijst, B.A.; Kusters, J.G.
J. Gen. Microbiol. 138, 2697-2706, 1992
A;Title: The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are composed c
A;Reference number: A47689; MUID:93139764; PMID:1487733
A;Contents: C5, Treponema
A;Accession: B47689
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-25 <KOO>
A;Cross-references: UNIPROT:P80158; UNIPARC:UPI000012A8D8
A;Note: sequence extracted from NCBI backbone (NCBIP:123401)

  Query Match             100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 4.3e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NAQRTLKF 18


RESULT 9
S14035
hypothetical protein - phage BZ13 (fragment)
C;Species: phage BZ13
C;Date: 18-Feb-1994 #sequence_revision 24-Apr-1998 #text_change 24-Apr-1998
C;Accession: S14035
R;Inokuchi, Y.; Hirashima, A.; Watanabe, I.
J. Mol. Biol. 158, 711-730, 1982
A;Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs from R
A;Reference number: S07250; MUID:83010313; PMID:7120417
A;Accession: S14035
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-26 <INO>
A;Cross-references: UNIPARC:UPI000017A891; EMBL:J02446; NID:g166167

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 GEPRSLYF 23


RESULT 10
S14036
hypothetical protein - phage GA (fragment)
C;Species: phage GA
C;Date: 18-Feb-1994 #sequence_revision 24-Apr-1998 #text_change 24-Apr-1998
C;Accession: S14036
R;Inokuchi, Y.; Hirashima, A.; Watanabe, I.
J. Mol. Biol. 158, 711-730, 1982
A;Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs from .
A;Reference number: S07250; MUID:83010313; PMID:7120417
A;Accession: S14036
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-26 <INO>
A;Cross-references: UNIPARC:UPI000017A895; EMBL:J02455; NID:g215437

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 GEPRSLYF 23


RESULT 11
S14037
hypothetical protein - phage JP34 (fragment)
C;Species: phage JP34
C;Date: 12-Feb-1998 #sequence_revision 17-Apr-1998 #text_change 17-Apr-1998
C;Accession: S14037
R;Inokuchi, Y.; Hirashima, A.; Watanabe, I.
J. Mol. Biol. 158, 711-730, 1982
A;Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs from .
A;Reference number: S07250; MUID:83010313; PMID:7120417
A;Accession: S14037
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-26 <INO>
A;Cross-references: UNIPARC:UPI000017A895; EMBL:J02456; NID:g215080

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 GEPRSLYF 23
```

```
RESULT 12
S13989
hypothetical protein - phage TH1 (fragment)
C;Species: phage TH1
C;Date: 18-Feb-1994 #sequence_revision 13-Mar-1998 #text_change 24-Apr-1998
C;Accession: S13989; S14039
R;Inokuchi, Y.; Hirashima, A.; Watanabe, I.
J. Mol. Biol. 158, 711-730, 1982
A;Title: Comparison of the nucleotide sequences at the 3'-terminal region of RNAs from R
A;Reference number: S07250; MUID:83010313; PMID:7120417
A;Accession: S13989
A;Status: translation not shown
A;Molecule type: genomic RNA
A;Residues: 1-26 <INO>
A;Cross-references: UNIPARC:UPI000017A89A; EMBL:J02519; NID:g216179

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 4.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 GELRSLYF 23


RESULT 13
S41774
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Trypanosoma congolense mit
C;Species: mitochondrion Trypanosoma congolense
C;Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 09-Jul-2004
C;Accession: S41774
R;Read, L.K.; Fish, W.R.; Muthiani, A.M.; Stuart, K.
Nucleic Acids Res. 21, 4073-4078, 1993
A;Title: Maxicircle DNA and edited mRNA sequences of closely related trypanosome species
A;Reference number: S41774; MUID:93382785; PMID:8396763
A;Accession: S41774
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-28 <REA>
A;Cross-references: UNIPROT:Q7M2D7; UNIPARC:UPI0000174C9E
C;Superfamily: cytochrome b; cytochrome b homology; cytochrome b6 homology; plastoquinol
C;Keywords: electron transfer; mitochondrion; oxidative phosphorylation; oxidoreductase;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 4.8e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 FRCRFLLF 9


RESULT 14
PT0366
T-cell receptor beta chain V-J region (6R2) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 30-May-1997
C;Accession: PT0366
R;Lehmann, P.V.; Drexler, K.; Tary-Lehmann, M.; Falcioni, F.; Hurtenbach, U.; Nagy, Z.A.
J. Exp. Med. 173, 333-341, 1991
A;Title: Graft-versus-host resistance induced by class II major histocompatibility compl
A;Reference number: PT0360; MUID:91108330; PMID:1824856
A;Accession: PT0366
A;Molecule type: mRNA
A;Residues: 1-28 <LEH>
A;Cross-references: UNIPARC:UPI000017C855
C;Keywords: T-cell receptor

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 4.8e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 FLERGLDF 23


RESULT 15
A44913
34K core flagella protein - Leptospira interrogans (fragment)
C;Species: Leptospira interrogans
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: A44913; A41210
R;Trueba, G.A.; Bolin, C.A.; Zuerner, R.L.
J. Bacteriol. 174, 4761-4768, 1992
A;Title: Characterization of the periplasmic flagellum proteins of Leptospira interrogan
A;Reference number: A44913; MUID:92325069; PMID:1624463
A;Accession: A44913
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-30 <TRU>
A;Cross-references: UNIPROT:Q9R5K3; UNIPARC:UPI00000BC595
A;Experimental source: sv. pomona type kennewicki
A;Note: sequence extracted from NCBI backbone (NCBIP:108220)
C;Superfamily: flagellin

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 5.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NSHRVLKF 18


RESULT 16
S74112
proline-rich antibacterial protein - green crab (fragment)
C;Species: Carcinus maenas (green crab, common shore crab)
C;Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 09-Jul-2004
C;Accession: S74112
R;Schnapp, D.; Kemp, G.D.; Smith, V.J.
Eur. J. Biochem. 240, 532-539, 1996
A;Title: Purification and characterization of a proline-rich antibacterial peptide, with
A;Reference number: S74112; MUID:97008941; PMID:8856051
A;Accession: S74112
A;Molecule type: protein
A;Residues: 1-30 <SCH>
A;Cross-references: UNIPROT:P82964; UNIPARC:UPI0000125BD2
A;Experimental source: haemocytes
C;Keywords: antibacterial

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 5.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         14 IGPRPLPF 21


RESULT 17
S53153
gene X protein - hepatitis B virus (isolate patient Usai'89) (fragment)
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Usai'89
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S53153
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53153
A;Molecule type: DNA
A;Residues: 1-31 <LAI>
A;Cross-references: UNIPROT:Q67974; UNIPARC:UPI00000F22D0; EMBL:X85270; NID:g736037; PID
```

```
A;Experimental source: isolate patient Usai'89
C;Genetics:
A;Gene: X
C;Superfamily: hepatitis B virus gene X protein

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 5.4e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 ADSRLLIF 9


RESULT 18
S53192
gene X protein - hepatitis B virus (isolate patient Italoc'92) (fragment)
C;Species: hepatitis B virus, HBV
A;Variety: isolate patient Italoc'92
C;Date: 08-Jul-1995 #sequence_revision 03-Aug-1995 #text_change 09-Jul-2004
C;Accession: S53192
R;Lai, M.E.; Mazzoleni, A.P.; Porru, A.; Balestrieri, A.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53112
A;Accession: S53192
A;Molecule type: DNA
A;Residues: 1-31 <LAI>
A;Cross-references: UNIPROT:Q68005; UNIPARC:UPI00000F9220; EMBL:X85257; NID:g736091; PID
A;Experimental source: isolate patient Italoc'92
C;Genetics:
A;Gene: X
C;Superfamily: hepatitis B virus gene X protein

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 5.4e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 EELRLLVF 9


RESULT 19
S39785
kallidin-releasing proteinase (EC 3.4.21.-) - puff adder (fragment)
C;Species: Bitis arietans (puff adder)
C;Date: 27-May-1994 #sequence_revision 19-Apr-1996 #text_change 09-Jul-2004
C;Accession: S39785
R;Nikai, T.; Momose, M.; Okumura, Y.; Ohara, A.; Komori, Y.; Sugihara, H.
Arch. Biochem. Biophys. 307, 304-310, 1993
A;Title: Kallidin-releasing enzyme from Bitis arietans (puff adder) venom.
A;Reference number: S39785; MUID:94099611; PMID:8274016
A;Accession: S39785
A;Molecule type: protein
A;Residues: 1-32 <NIK>
A;Cross-references: UNIPROT:Q9PRY9; UNIPARC:UPI00000FD0E5
C;Superfamily: trypsin; trypsin homology
C;Keywords: glycoprotein; hydrolase; serine proteinase; venom

  Query Match             100.0%;  Score 20;  DB 2;  Length 32;
  Best Local Similarity   37.5%;  Pred. No. 5.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         10 NEHRFLVF 17


RESULT 20
B82421
hypothetical protein VCA0761 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: B82421
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: B82421
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-32 <HEI>
A;Cross-references: UNIPROT:Q9KLI3; UNIPARC:UPI00000C3636; GB:AE004404; GB:AE003853; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0761
A;Map position: 2

  Query Match             100.0%;  Score 20;  DB 2;  Length 32;
  Best Local Similarity   37.5%;  Pred. No. 5.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MPDRLLRF 8


RESULT 21
F81044
hypothetical protein NMB1778 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81044
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.P
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: F81044
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-34 <TET>
A;Cross-references: UNIPROT:Q9JY24; UNIPARC:UPI00000C47BF; GB:AE002527; GB:AE002098; NI
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1778

  Query Match             100.0%;  Score 20;  DB 2;  Length 34;
  Best Local Similarity   37.5%;  Pred. No. 5.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          4 KKLRFLIF 11


RESULT 22
H83722
hypothetical protein BH0584 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83722
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83722
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-34 <STO>
A;Cross-references: UNIPROT:Q9KFA1; UNIPARC:UPI00000C390F; GB:AP001509; GB:BA000004; NI
A;Experimental source: strain C-125
```

```
C;Genetics:
A;Gene: BH0584

  Query Match             100.0%;  Score 20;  DB 2;  Length 34;
  Best Local Similarity   37.5%;  Pred. No. 5.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MKKRSLVF 8


RESULT 23
A81164
hypothetical protein NMB0754 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: A81164
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81164
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-36 <TET>
A;Cross-references: UNIPROT:Q9K066; UNIPARC:UPI00000C4566; GB:AE002429; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0754

  Query Match             100.0%;  Score 20;  DB 2;  Length 36;
  Best Local Similarity   37.5%;  Pred. No. 6.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         13 SNQRSLRF 20


RESULT 24
A69326
hypothetical protein AF0609 - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: A69326
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: A69326
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-36 <KLE>
A;Cross-references: UNIPROT:O29646; UNIPARC:UPI0000057051; GB:AE001063; GB:AE000782; NID

  Query Match             100.0%;  Score 20;  DB 2;  Length 36;
  Best Local Similarity   37.5%;  Pred. No. 6.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         28 VRIRNLQF 35


RESULT 25
A59185
photosystem II protein psbL - Prochlorothrix hollandica
C;Species: Prochlorothrix hollandica
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C;Accession: A59185
R;Bullerjahn, G.
submitted to the Protein Sequence Database, February 2000
A;Reference number: A59185
A;Accession: A59185
A;Molecule type: protein
A;Residues: 1-38 <BUL>
A;Cross-references: UNIPROT:Q7M157; UNIPARC:UPI0000178142
A;Experimental source: strain ACC15-2
A;Note: 4.5K transmembrane component of photosystem II
C;Superfamily: photosystem II protein psbL
C;Keywords: photosynthesis; photosystem II; transmembrane protein
F;18-37/Domain: transmembrane #status predicted <TMM>

  Query Match             100.0%;  Score 20;  DB 2;  Length 38;
  Best Local Similarity   37.5%;  Pred. No. 6.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         19 FLGRLLIF 26


RESULT 26
D84227
hypothetical protein Vng0697h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84227
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84227
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <STO>
A;Cross-references: UNIPROT:Q9HRH4; UNIPARC:UPI00000636DF; GB:AE004437; NID:g10580280; P
C;Genetics:
A;Gene: VNG0697H

  Query Match             100.0%;  Score 20;  DB 2;  Length 38;
  Best Local Similarity   37.5%;  Pred. No. 6.5e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         22 PDGRTLEF 29


RESULT 27
H81579
hypothetical protein CP0403 [imported] - Chlamydophila pneumoniae (strain AR39)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: H81579
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: H81579
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-38 <REA>
A;Cross-references: UNIPROT:Q9K283; UNIPARC:UPI00000CCC62; GB:AE002202; GB:AE002161; NID
```

A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0403

Query Match 100.0%; Score 20; DB 2; Length 38;
Best Local Similarity 37.5%; Pred. No. 6.5e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 22 KGKRELFF 29

RESULT 28
A81151
hypothetical protein NMB0862 [imported] - Neisseria meningitidis (strain MC58 serogroup
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: A81151
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755; PMID:10710307
A;Accession: A81151
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <TET>
A;Cross-references: UNIPROT:Q9JZX2; UNIPARC:UPI00000C45B1; GB:AE002439; GB:AE002098; NID
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB0862

Query Match 100.0%; Score 20; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 6.7e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 23 VLSRALIF 30

RESULT 29
B85990
hypothetical protein Z4614 [imported] - Escherichia coli (strain O157:H7, substrain EDL9
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C;Accession: B85990
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: B85990
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-39 <STO>
A;Cross-references: UNIPROT:Q8X417; UNIPARC:UPI00000D0E65; GB:AE005174; NID:g12517881; P
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: Z4614

Query Match 100.0%; Score 20; DB 2; Length 39;
Best Local Similarity 37.5%; Pred. No. 6.7e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 25 LILRALAF 32

RESULT 30
S07969
T-cell receptor alpha chain V-J region (clone 18 BBM 7) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 23-Jul-1999
C;Accession: S07969
R;Bill, J.; Yaguee, J.; Appel, V.B.; White, J.; Horn, G.; Erlich, H.A.; Palmer, E.
J. Exp. Med. 169, 115-133, 1989
A;Title: Molecular genetic analysis of 178 I-A(bm12)-reactive T cells.
A;Reference number: S05590; MUID:89080476; PMID:2783331
A;Accession: S07969
A;Molecule type: mRNA
A;Residues: 1-40 <BIL>
A;Cross-references: UNIPARC:UPI0000115E1A; EMBL:X14931; NID:g54859; PIDN:CAA33058.1; PID
A;Note: the protein sequence from Fig. 4B is inconsistent with the nucleotide sequence f
A;Note: this sequence was determined from the differentiated gene
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor
F;1-15/Region: V segment (fragment)
F;17-35/Region: J segment

Query Match 100.0%; Score 20; DB 2; Length 40;
Best Local Similarity 37.5%; Pred. No. 6.9e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 18 QGGRALIF 25

RESULT 31
B27579
T-cell receptor beta chain V-D-J regions (TRB67) - rat (fragment)
C;Species: Rattus norvegicus (Norway rat)
C;Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 23-Jul-1999
C;Accession: B27579
R;Morris, M.; Barclay, A.N.; Williams, A.F.
Immunogenetics 27, 174-179, 1988
A;Title: Analysis of T cell receptor beta chains in rat thymus, and rat C-alpha and C-b
A;Reference number: A27578; MUID:88113841; PMID:2962935
A;Accession: B27579
A;Molecule type: mRNA
A;Residues: 1-41 <MOR>
A;Cross-references: UNIPARC:UPI0000114D4B; GB:M18849; NID:g207241; PIDN:AAA42229.1; PID
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: T-cell receptor

Query Match 100.0%; Score 20; DB 2; Length 41;
Best Local Similarity 37.5%; Pred. No. 7.1e+02;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 12 PDDRGLYF 19

RESULT 32
D82691
hypothetical protein XF1349 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: D82691
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seque
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: D82691
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <SIM>
A;Cross-references: UNIPROT:Q9PDN0; UNIPARC:UPI00000C26D8; GB:AE003967; GB:AE003849; NI
A;Experimental source: strain 9a5c

R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1349

  Query Match 100.0%; Score 20; DB 2; Length 41;
  Best Local Similarity 37.5%; Pred. No. 7.1e+02;
  Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
  :::|:|:|
Db 32 MTDRNLIF 39

RESULT 33
E82544
hypothetical protein XF2541 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82544
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82544
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-41 <SIM>
A;Cross-references: UNIPROT:Q9PAH7; UNIPARC:UPI00000C2A90; GB:AE004061; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF2541

  Query Match 100.0%; Score 20; DB 2; Length 41;
  Best Local Similarity 37.5%; Pred. No. 7.1e+02;
  Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
  :::|:|:|
Db 20 WCQRYLGF 27

RESULT 34
G56271
lpfA 5'-region hypothetical protein - Salmonella typhimurium (fragment)
C;Species: Salmonella typhimurium
C;Date: 03-Oct-1995 #sequence_revision 03-Oct-1995 #text_change 29-Sep-1999
C;Accession: G56271
R;Baeumler, A.J.; Heffron, F.
J. Bacteriol. 177, 2087-2097, 1995
A;Title: Identification and sequence analysis of lpfABCDE, a putative fimbrial operon of
A;Reference number: A56271; MUID:95238281; PMID:7721701
A;Accession: G56271
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-42 <BAE>
A;Cross-references: UNIPARC:UPI00001702D6; GB:U18559; NID:g829370; PIDN:AAA73965.1; PID:
C;Superfamily: hypothetical protein c0103

  Query Match 100.0%; Score 20; DB 2; Length 42;
  Best Local Similarity 37.5%; Pred. No. 7.2e+02;
  Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
  :::|:|:|
Db 5 GGSRVLTF 12

RESULT 35
T07581
hypothetical protein 42g - Japanese black pine chloroplast
C;Species: chloroplast Pinus thunbergiana (Japanese black pine)
C;Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 09-Jul-2004
C;Accession: T07581
R;Wakasugi, T.; Tsudzuki, J.; Ito, S.; Nakashima, K.; Tsudzuki, T.; Sugiura, M.
Proc. Natl. Acad. Sci. U.S.A. 91, 9794-9798, 1994
A;Title: Loss of all ndh genes as determined by sequencing the entire chloroplast genome
A;Reference number: Z16030; MUID:95024047; PMID:7937893
A;Accession: T07581
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-42 <WAK>
A;Cross-references: UNIPROT:Q33005; UNIPARC:UPI00000939E4; EMBL:D17510; NID:g529643; PID
C;Genetics:
A;Genome: chloroplast
C;Keywords: chloroplast

  Query Match 100.0%; Score 20; DB 2; Length 42;
  Best Local Similarity 37.5%; Pred. No. 7.2e+02;
  Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
  :::|:|:|
Db 11 LSIRSLSF 18

RESULT 36
D81730
hypothetical protein TC0191 [imported] - Chlamydia muridarum (strain Nigg)
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81730
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: D81730
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-42 <TET>
A;Cross-references: UNIPROT:Q9PLB6; UNIPARC:UPI000005781F; GB:AE002286; GB:AE002160; NID
A;Experimental source: strain Nigg (MoPn)
C;Genetics:
A;Gene: TC0191

  Query Match 100.0%; Score 20; DB 2; Length 42;
  Best Local Similarity 37.5%; Pred. No. 7.2e+02;
  Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XXXRXLXF 8
              :::|:|:|
Db         30 IGVRQLEF 37
```

RESULT 37
B82629
hypothetical protein XF1866 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: B82629
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-42 <SIM>
A;Cross-references: UNIPROT:Q9PCB5; UNIPARC:UPI00000C286C; GB:AE004007; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laign
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF1866

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 42;
  Best Local Similarity   37.5%;  Pred. No. 7.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MYVRQLDF 8
```

RESULT 38
F81505
hypothetical protein CP1078 [imported] - Chlamydophila pneumoniae (strain AR39)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: F81505
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: F81505
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-43 <REA>
A;Cross-references: UNIPROT:Q9K1S7; UNIPARC:UPI00000CCCE0; GB:AE002264; GB:AE002161; NID
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP1078

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 43;
  Best Local Similarity   37.5%;  Pred. No. 7.4e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         10 SLRRHLLF 17
```

RESULT 39
PL0091
Ig heavy chain V region (B3938) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 23-Jul-1999
C;Accession: PL0091
R;Meek, K.; Hasemann, C.; Pollok, B.; Alkan, S.S.; Brait, M.; Slaoui, M.; Urbain, J.; Ca
J. Exp. Med. 169, 519-533, 1989
A;Title: Structural characterization of antiidiotypic antibodies; evidence that Ab2s are
A;Reference number: PL0080; MUID:89094248; PMID:2492056
A;Accession: PL0091
A;Molecule type: mRNA
A;Residues: 1-44 <MEE>
A;Cross-references: UNIPARC:UPI0000115F25; GB:X58591; GB:Y00794; NID:g51567; PIDN:CAA414
A;Note: the sequence shown here is from the VH region of a syngeneic antibody to antiphc
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 44;
  Best Local Similarity   37.5%;  Pred. No. 7.6e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         24 PLLRLLHF 31
```

RESULT 40
H83936
hypothetical protein BH2296 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H83936
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H83936
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-45 <STO>
A;Cross-references: UNIPROT:Q9KAJ1; UNIPARC:UPI00000C3E80; GB:AP001515; GB:BA000004; NIE
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH2296

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 45;
  Best Local Similarity   37.5%;  Pred. No. 7.7e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 RKIRKLSF 9
```

RESULT 41
F95023
hypothetical protein SP0203 [imported] - Streptococcus pneumoniae (strain TIGR4)
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95023
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Hei
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple,
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916

```
A;Accession: F95023
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <KUR>
A;Cross-references: UNIPROT:Q97SW1; UNIPARC:UPI000005133F; GB:AE005672; PIDN:AAK74383.1;
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0203

  Query Match             100.0%;  Score 20;  DB 2;  Length 46;
  Best Local Similarity   37.5%;  Pred. No. 7.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         28 KKERLLSF 35


RESULT 42
A99802
hypothetical protein ECs1385 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: A99802
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99802
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <HAY>
A;Cross-references: UNIPROT:Q8X2G6; UNIPARC:UPI00000D2AD3; GB:BA000007; PIDN:BAB34808.1;
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1385

  Query Match             100.0%;  Score 20;  DB 2;  Length 46;
  Best Local Similarity   37.5%;  Pred. No. 7.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         12 HGHRLLCF 19


RESULT 43
D84334
hypothetical protein Vng1832h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: D84334
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: D84334
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <STO>
A;Cross-references: UNIPROT:Q9HP30; UNIPARC:UPI000006399F; GB:AE004437; NID:g10581278; P
C;Genetics:
A;Gene: VNG1832H

  Query Match             100.0%;  Score 20;  DB 2;  Length 46;
  Best Local Similarity   37.5%;  Pred. No. 7.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          4 NEKRRLLF 11


RESULT 44
AG3569
hypothetical protein BMEII0480 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 09-Jul-2004
C;Accession: AG3569
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
A;Reference number: AD3252; PMID:11756688
A;Accession: AG3569
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-46 <KUR>
A;Cross-references: UNIPROT:Q8YCP8; UNIPARC:UPI00000584CF; GB:AE008918; PIDN:AAL53722.1;
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEII0480
A;Map position: II

  Query Match             100.0%;  Score 20;  DB 2;  Length 46;
  Best Local Similarity   37.5%;  Pred. No. 7.9e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         33 SGKRSLKF 40


RESULT 45
S39358
cyclin kinase inhibitor - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 25-Feb-1994 #sequence_revision 17-Nov-1995 #text_change 17-Mar-1999
C;Accession: S39358
R;Xiong, Y.; Hannon, G.J.; Zhang, H.; Casso, D.; Kobayashi, R.; Beach, D.
Nature 366, 701-704, 1993
A;Title: p21 is a universal inhibitor of cyclin kinases.
A;Reference number: S39357; MUID:94081955; PMID:8259214
A;Accession: S39358
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-47 <XIO>
A;Cross-references: UNIPARC:UPI0000179862

  Query Match             100.0%;  Score 20;  DB 2;  Length 47;
  Best Local Similarity   37.5%;  Pred. No. 8.1e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         38 HSKRRLIF 45


RESULT 46
T20751
hypothetical protein F11A5.6 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T20751
R;Gardner, A.
submitted to the EMBL Data Library, March 1997
A;Reference number: Z19319
A;Accession: T20751
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-48 <WIL>
A;Cross-references: UNIPROT:O17790; UNIPARC:UPI0000078C38; EMBL:Z92830; PIDN:CAB07358.1,
```

```
A;Experimental source: clone F11A5
C;Genetics:
A;Gene: CESP:F11A5.6
A;Map position: 5

  Query Match             100.0%;  Score 20;  DB 2;  Length 48;
  Best Local Similarity   37.5%;  Pred. No. 8.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         19 EVMRILGF 26


RESULT 47
H84063
hypothetical protein BH3312 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: H84063
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: H84063
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-48 <STO>
A;Cross-references: UNIPROT:Q9K7P9; UNIPARC:UPI00000C416D; GB:AP001518; GB:BA000004; NID
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3312

  Query Match             100.0%;  Score 20;  DB 2;  Length 48;
  Best Local Similarity   37.5%;  Pred. No. 8.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         33 MVRRCLPF 40


RESULT 48
B28118
ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b - Crithidia fasciculata mitc
C;Species: mitochondrion Crithidia fasciculata
C;Date: 28-Oct-1988 #sequence_revision 28-Oct-1988 #text_change 31-Dec-2004
C;Accession: B28118; B25877
R;Feagin, J.E.; Shaw, J.M.; Simpson, L.; Stuart, K.
Proc. Natl. Acad. Sci. U.S.A. 85, 539-543, 1988
A;Title: Creation of AUG initiation codons by addition of uridines within cytochrome b t
A;Reference number: A28118; MUID:88124876; PMID:2448777
A;Accession: B28118
A;Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A;Molecule type: mRNA
A;Residues: 1-48 <FEA>
A;Cross-references: UNIPROT:Q34098; UNIPARC:UPI0000174C9D
R;Sloof, P.; van den Burg, J.; Voogd, A.; Benne, R.
Nucleic Acids Res. 15, 51-65, 1987
A;Title: The nucleotide sequence of a 3.2 kb segment of mitochondrial maxicircle DNA frc
ytochrome b gene and a possible frameshift gene; further evidence for the use of unusual
A;Reference number: A25877; MUID:87146364; PMID:3029678
A;Accession: B25877
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 'KR',15,'KKKR',20-48,'W' <SLO>
A;Cross-references: UNIPARC:UPI000009677F; GB:X05063; NID:g12861; PIDN:CAA28730.1; PID:g
C;Genetics:
A;Genome: mitochondrion
A;Genetic code: SGC6
C;Superfamily: cytochrome b homology; cytochrome b6 homology; plastoquinol-plastocyanin
C;Keywords: electron transfer; mitochondrion; oxidative phosphorylation; oxidoreductase;

  Query Match             100.0%;  Score 20;  DB 2;  Length 48;
  Best Local Similarity   37.5%;  Pred. No. 8.2e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          3 FRVRFLLF 10


RESULT 49
D81567
hypothetical protein CP0514 [imported] - Chlamydophila pneumoniae (strain AR39)
C;Species: Chlamydophila pneumoniae, Chlamydia pneumoniae
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: D81567
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
, C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: A81500; MUID:20150255; PMID:10684935
A;Accession: D81567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-49 <REA>
A;Cross-references: UNIPROT:Q9K257; UNIPARC:UPI00000CCC76; GB:AE002212; GB:AE002161; NIL
A;Experimental source: strain AR39, HL cells
C;Genetics:
A;Gene: CP0514

  Query Match             100.0%;  Score 20;  DB 2;  Length 49;
  Best Local Similarity   37.5%;  Pred. No. 8.4e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         33 FKLRDLLF 40


RESULT 50
H90537
hypothetical protein MYPU_2080 [imported] - Mycoplasma pulmonis (strain UAB CTIP)
C;Species: Mycoplasma pulmonis
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: H90537
R;Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.,
Nucleic Acids Res. 29, 2145-2153, 2001
A;Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pul
A;Reference number: A99512; MUID:21267165; PMID:11353084
A;Accession: H90537
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-50 <KUR>
A;Cross-references: UNIPROT:Q98R02; UNIPARC:UPI00000C802A; GB:AL445566; PID:g14089621;
A;Experimental source: strain UAB CTIP
C;Genetics:
A;Gene: MYPU_2080
A;Genetic code: SGC3

  Query Match             100.0%;  Score 20;  DB 2;  Length 50;
  Best Local Similarity   37.5%;  Pred. No. 8.6e+02;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         14 LTSRFLIF 21


Search completed: May  5, 2006, 12:23:53
Job time : 20 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:19:10 ; Search time 74 Seconds
(without alignments)
76.273 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 20
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	9	1	FAR9_ASCSU	P43172 ascaris suu
2	20	100.0	13	2	Q841R8_SYNP6	Q841r8 synechococc
3	20	100.0	17	2	Q64973_9BROM	Q64973 alfalfa mos
4	20	100.0	17	2	Q64974_9BROM	Q64974 alfalfa mos
5	20	100.0	19	2	Q711Y2_RHILO	Q711y2 rhizobium l
6	20	100.0	20	2	Q5C7I1_SCHJA	Q5c7i1 schistosoma
7	20	100.0	20	2	Q4YCL5_PLABE	Q4ycl5 plasmodium
8	20	100.0	20	2	Q711Y6_RHILO	Q711y6 rhizobium l
9	20	100.0	20	2	Q9R4W7_HELPY	Q9r4w7 helicobacte
10	20	100.0	20	2	Q9R5K2_LEPIN	Q9r5k2 leptospira
11	20	100.0	21	2	Q8XPJ2_RALSO	Q8xpj2 ralstonia s
12	20	100.0	22	2	Q33573_9TRYP	Q33573 trypanosoma
13	20	100.0	22	2	Q4YTG6_PLABE	Q4ytg6 plasmodium
14	20	100.0	22	2	Q9R510_HELPY	Q9r510 helicobacte
15	20	100.0	22	2	Q9K8Q3_BACHD	Q9k8q3 bacillus ha
16	20	100.0	23	2	Q5BR34_SCHJA	Q5br34 schistosoma
17	20	100.0	24	2	Q4X467_PLACH	Q4x467 plasmodium
18	20	100.0	24	2	Q7M132_TREHY	Q7m132 treponema h
19	20	100.0	24	2	Q4TJY7_9SPHN	Q4tjy7 erythrobact
20	20	100.0	25	1	FLAB1_TREHY	P80158 treponema h
21	20	100.0	25	2	Q5C711_SCHJA	Q5c711 schistosoma
22	20	100.0	25	2	Q4X8A2_PLACH	Q4x8a2 plasmodium
23	20	100.0	25	2	O11472_9HEPC	O11472 hepatitis c
24	20	100.0	25	2	Q9PRS4_ONCMY	Q9prs4 oncorhynchu
25	20	100.0	26	2	Q4XIF3_PLACH	Q4xif3 plasmodium
26	20	100.0	26	2	Q4XSJ0_PLACH	Q4xsj0 plasmodium
27	20	100.0	26	2	Q91ZN8_MOUSE	Q91zn8 mus musculu
28	20	100.0	27	2	Q7ZW70_BRARE	Q7zw70 brachydanio
29	20	100.0	27	2	Q4SKW5_TETNG	Q4skw5 tetraodon n
30	20	100.0	28	2	Q7M2D7_TRYCO	Q7m2d7 trypanosoma
31	20	100.0	28	2	Q4YJN9_PLABE	Q4yjn9 plasmodium
32	20	100.0	28	2	Q4L621_STAHJ	Q4l621 staphylococ
33	20	100.0	28	2	Q9DD70_CHICK	Q9dd70 gallus gall
34	20	100.0	28	2	Q9DFT7_CHICK	Q9dft7 gallus gall
35	20	100.0	30	1	AP65_CARMA	P82964 carcinus ma
36	20	100.0	30	2	Q4XCS4_PLACH	Q4xcs4 plasmodium
37	20	100.0	30	2	Q4YQ95_PLABE	Q4yq95 plasmodium
38	20	100.0	30	2	Q9R5K3_LEPIN	Q9r5k3 leptospira
39	20	100.0	30	2	Q9RER6_ENTAE	Q9rer6 enterobacte
40	20	100.0	30	2	Q57H84_SALCH	Q57h84 salmonella
41	20	100.0	30	2	Q5PFE0_SALPA	Q5pfe0 salmonella
42	20	100.0	31	2	Q4XBW6_PLACH	Q4xbw6 plasmodium
43	20	100.0	31	2	Q4KA34_PSEF5	Q4ka34 pseudomonas
44	20	100.0	31	2	Q65TM8_MANSM	Q65tm8 mannheimia
45	20	100.0	31	2	Q87K91_VIBPA	Q87k91 vibrio para
46	20	100.0	31	2	Q8KBJ8_CHLTE	Q8kbj8 chlorobium
47	20	100.0	31	2	Q67974_HPBV0	Q67974 hepatitis b
48	20	100.0	31	2	Q68005_HPBV0	Q68005 hepatitis b
49	20	100.0	32	2	Q9BYF3_HUMAN	Q9byf3 homo sapien
50	20	100.0	32	2	Q4X2P5_PLACH	Q4x2p5 plasmodium
51	20	100.0	32	2	Q7P2W5_FUSNV	Q7p2w5 fusobacteri
52	20	100.0	32	2	Q65SZ4_MANSM	Q65sz4 mannheimia
53	20	100.0	32	2	Q9KLI3_VIBCH	Q9kli3 vibrio chol
54	20	100.0	32	2	Q81S50_BACAN	Q81s50 bacillus an
55	20	100.0	32	2	Q5HGK1_STAAC	Q5hgk1 staphylococ
56	20	100.0	32	2	Q8KF24_CHLTE	Q8kf24 chlorobium
57	20	100.0	32	2	Q69435_9FLAV	Q69435 gb virus c/
58	20	100.0	32	2	Q69436_9FLAV	Q69436 gb virus c/
59	20	100.0	32	2	Q69438_9FLAV	Q69438 gb virus c/
60	20	100.0	32	2	Q69439_9FLAV	Q69439 gb virus c/
61	20	100.0	32	2	Q69440_9FLAV	Q69440 gb virus c/
62	20	100.0	32	2	Q69441_9FLAV	Q69441 gb virus c/
63	20	100.0	32	2	Q69442_9FLAV	Q69442 gb virus c/
64	20	100.0	32	2	Q9PXV2_HPBV0	Q9pxv2 hepatitis b
65	20	100.0	32	2	O13040_LAMPL	O13040 lampetra pl
66	20	100.0	32	2	Q9PRY9_BITAR	Q9pry9 bitis ariet
67	20	100.0	33	2	Q9BUZ1_HUMAN	Q9buz1 homo sapien
68	20	100.0	33	2	Q5BY27_SCHJA	Q5by27 schistosoma
69	20	100.0	33	2	Q4Z9N3_9VIRU	Q4z9n3 bacteriopha
70	20	100.0	33	2	Q8GQU2_LEPBO	Q8gqu2 leptospira
71	20	100.0	33	2	Q93Q13_LISMO	Q93q13 listeria mo
72	20	100.0	33	2	Q9R5L6_HELPY	Q9r516 helicobacte
73	20	100.0	33	2	Q4MTJ7_BACCE	Q4mtj7 bacillus ce
74	20	100.0	33	2	Q731V9_BACC1	Q731v9 bacillus ce
75	20	100.0	33	2	Q74NL2_BACC1	Q74nl2 bacillus ce
76	20	100.0	33	2	Q8EJ76_SHEON	Q8ej76 shewanella
77	20	100.0	33	2	Q81MM6_BACAN	Q81mm6 bacillus an
78	20	100.0	33	2	O11474_9FLAV	O11474 gb virus c/
79	20	100.0	33	2	O11475_9FLAV	O11475 gb virus c/
80	20	100.0	33	2	O11476_9FLAV	O11476 gb virus c/
81	20	100.0	33	2	O11477_9FLAV	O11477 gb virus c/
82	20	100.0	33	2	O39192_9FLAV	O39192 gb virus c/
83	20	100.0	33	2	O39193_9FLAV	O39193 gb virus c/
84	20	100.0	33	2	O39194_9FLAV	O39194 gb virus c/
85	20	100.0	33	2	Q5EGW5_ANAPA	Q5egw5 anas falcat
86	20	100.0	33	2	Q4RLD7_TETNG	Q4rld7 tetraodon n
87	20	100.0	34	2	Q8WW51_HUMAN	Q8ww51 homo sapien
88	20	100.0	34	2	Q4XTE8_PLACH	Q4xte8 plasmodium
89	20	100.0	34	2	Q43108_PTEVI	Q43108 pteris vitt
90	20	100.0	34	2	Q5QET2_ARAHY	Q5qet2 arachis hyp
91	20	100.0	34	2	Q8F1K5_LEPIN	Q8f1k5 leptospira
92	20	100.0	34	2	Q87RX1_VIBPA	Q87rx1 vibrio para
93	20	100.0	34	2	Q9JY24_NEIMB	Q9jy24 neisseria m
94	20	100.0	34	2	Q9KFA1_BACHD	Q9kfa1 bacillus ha
95	20	100.0	34	2	Q729W1_DESVH	Q729w1 desulfovibr
96	20	100.0	34	2	Q5SU80_MOUSE	Q5su80 mus musculu
97	20	100.0	34	2	Q4RCN8_TETNG	Q4rcn8 tetraodon n
98	20	100.0	34	2	Q4TH65_TETNG	Q4th65 tetraodon n
99	20	100.0	35	2	Q7S6V3_NEUCR	Q7s6v3 neurospora
100	20	100.0	35	2	Q8ITS8_GALME	Q8its8 galleria me

ALIGNMENTS

```
RESULT 1
FAR9_ASCSU
ID   FAR9_ASCSU     STANDARD;      PRT;     9 AA.
AC   P43172;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide AF9.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=95380362; PubMed=7651904; DOI=10.1016/0196-9781(94)00211-N;
RA   Cowden C., Stretton A.O.W.;
RT   "Eight novel FMRFamide-like neuropeptides isolated from the nematode
RT   Ascaris suum.";
RL   Peptides 16:491-500(1995).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       9      9       Phenylalanine amide.
SQ   SEQUENCE   9 AA;  1012 MW;  524F073774176877 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 9;
  Best Local Similarity   37.5%;  Pred. No. 2.2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 LGPRPLRF 9


RESULT 2
Q841R8_SYNP6
ID   Q841R8_SYNP6 PRELIMINARY;      PRT;    13 AA.
AC   Q841R8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   CpcH (Fragment).
GN   Name=cpcH;
OS   Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=269084;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93231540; PubMed=7682531; DOI=10.1016/0378-1119(93)90592-Q;
RA   Kalla R., Bhalerao R.P., Gustafsson P.;
RT   "Regulation of phycobilisome rod proteins and mRNA at different light
RT   intensities in the cyanobacterium Synechococcus 6301.";
RL   Gene 126:77-83(1993).
DR   EMBL; S58974; AAP13908.1; -; Genomic_DNA.
FT   NON_TER      13     13
SQ   SEQUENCE   13 AA;  1439 MW;  D5F6F2E86F2D0DDD CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 13;
  Best Local Similarity   37.5%;  Pred. No. 1.2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          5 EAARRLGF 12


RESULT 3
Q64973_9BROM
ID   Q64973_9BROM PRELIMINARY;      PRT;    17 AA.
AC   Q64973;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   (strain AlMV-S) 5' end of RNA-2. (Fragment).
OS   Alfalfa mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC   Alfamovirus.
OX   NCBI_TaxID=12321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AlMV-S;
RX   MEDLINE=83220821; PubMed=6856476;
RA   Ravelonandro M., Godefroy-Colburn T., Pinck L.;
RT   "Structure of the 5'-terminal untranslated region of the genomic RNAs
RT   from two strains of alfalfa mosaic virus.";
RL   Nucleic Acids Res. 11:2815-2826(1983).
DR   EMBL; M36389; AAA66595.1; -; Genomic_RNA.
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1915 MW;  0AE5B9E26070F42B CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 17;
  Best Local Similarity   37.5%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          3 TLLRCLGF 10


RESULT 4
Q64974_9BROM
ID   Q64974_9BROM PRELIMINARY;      PRT;    17 AA.
AC   Q64974;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   (strain AlMV-B) 5' end of RNA-2. (Fragment).
OS   Alfalfa mosaic virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Bromoviridae;
OC   Alfamovirus.
OX   NCBI_TaxID=12321;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AlMV-B;
RX   MEDLINE=83220821; PubMed=6856476;
RA   Ravelonandro M., Godefroy-Colburn T., Pinck L.;
RT   "Structure of the 5'-terminal untranslated region of the genomic RNAs
RT   from two strains of alfalfa mosaic virus.";
RL   Nucleic Acids Res. 11:2815-2826(1983).
DR   EMBL; M36390; AAA66596.1; -; Genomic_RNA.
FT   NON_TER      17     17
SQ   SEQUENCE   17 AA;  1915 MW;  0AE5B9E26070F42B CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 17;
  Best Local Similarity   37.5%;  Pred. No. 1.5e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          3 TLLRCLGF 10


RESULT 5
Q711Y2_RHILO
ID   Q711Y2_RHILO PRELIMINARY;      PRT;    19 AA.
AC   Q711Y2;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
```

```
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   NodB protein (Fragment).
GN   Name=nodB;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=NZP2037;
RA   Moulin L.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ300246; CAC82867.1; -; Genomic_DNA.
FT   NON_TER      19     19
SQ   SEQUENCE   19 AA;  2200 MW;  4FCBFE6848808BEB CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 19;
  Best Local Similarity   37.5%;  Pred. No. 1.7e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          6 LSARFLKF 13


RESULT 6
Q5C7I1_SCHJA
ID   Q5C7I1_SCHJA PRELIMINARY;      PRT;    20 AA.
AC   Q5C7I1;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY808504; AAX24393.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   20 AA;  2421 MW;  B0DABA3920860C4B CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 LESRLLRF 18


RESULT 7
Q4YCL5_PLABE
ID   Q4YCL5_PLABE PRELIMINARY;      PRT;    20 AA.
AC   Q4YCL5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=PB102933.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01006614; CAI04275.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   20 AA;  2470 MW;  DB7AA80E3BC317C4 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          5 VILRFLIF 12


RESULT 8
Q711Y6_RHILO
ID   Q711Y6_RHILO PRELIMINARY;      PRT;    20 AA.
AC   Q711Y6;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   NodB protein (Fragment).
GN   Name=nodB;
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=E1R;
RA   Moulin L.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ300244; CAC82863.1; -; Genomic_DNA.
FT   NON_TER      20     20
SQ   SEQUENCE   20 AA;  2274 MW;  F8CFDB240848808B CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          6 LSARFLKF 13


RESULT 9
Q9R4W7_HELPY
ID   Q9R4W7_HELPY  PRELIMINARY;      PRT;    20 AA.
AC   Q9R4W7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   66 kDa major heat shock protein (Fragment).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=95020803; PubMed=7935068;
RA   Yokota K., Hirai Y., Haque M., Hayashi S., Isogai H., Sugiyama T.,
RA   Nagamachi E., Tsukada Y., Fujii N., Oguma K.;
RT   "Heat shock protein produced by Helicobacter pylori.";
RL   Microbiol. Immunol. 38:403-405(1994).
SQ   SEQUENCE   20 AA;  2368 MW;  D5C93C2B277B4EA1 CRC64;
```

```
  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          8 DSARNLLF 15


RESULT 10
Q9R5K2_LEPIN
ID   Q9R5K2_LEPIN  PRELIMINARY;      PRT;    20 AA.
AC   Q9R5K2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   35.5 kDa periplasmic flagella core protein (Fragment).
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92325069; PubMed=1624463;
RA   Trueba G.A., Bolin C.A., Zuerner R.L.;
RT   "Characterization of the periplasmic flagellum proteins of Leptospira
RT   interrogans.";
RL   J. Bacteriol. 174:4761-4768(1992).
DR   PIR; B44913; B44913.
SQ   SEQUENCE   20 AA;  2328 MW;  D6C2C8EC3584C5AC CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 20;
  Best Local Similarity   37.5%;  Pred. No. 1.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 FAHRTLKF 18


RESULT 11
Q8XPJ2_RALSO
ID   Q8XPJ2_RALSO PRELIMINARY;      PRT;    21 AA.
AC   Q8XPJ2;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein RSp1648.
GN   OrderedLocusNames=RSp1648; ORFNames=RS02207;
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OG   Plasmid megaplasmid.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Burkholderiaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852; DOI=10.1038/415497a;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646086; CAD18799.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein; Plasmid.
SQ   SEQUENCE   21 AA;  2191 MW;  D33004D2CFBE3869 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 21;
  Best Local Similarity   37.5%;  Pred. No. 1.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          7 DGARVLLF 14


RESULT 12
Q33573_9TRYP
ID   Q33573_9TRYP PRELIMINARY;      PRT;    22 AA.
AC   Q33573;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome b (Fragment).
OS   Trypanosoma brucei.
OG   Mitochondrion.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5691;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=EATRO 164;
RX   MEDLINE=88124876; PubMed=2448777;
RA   Feagin J.E., Shaw J.M., Simpson L., Stuart K.;
RT   "Creation of AUG initiation codons by addition of uridines within
RT   cytochrome b transcripts of kinetoplastids.";
RL   Proc. Natl. Acad. Sci. U.S.A. 85:539-543(1988).
DR   EMBL; M19064; AAA32116.1; -; mRNA.
FT   NON_TER      22     22
SQ   SEQUENCE   22 AA;  2788 MW;  3C559E3C3BAFA636 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 22;
  Best Local Similarity   37.5%;  Pred. No. 2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 FRCRFLLF 9


RESULT 13
Q4YTG6_PLABE
ID   Q4YTG6_PLABE PRELIMINARY;      PRT;    22 AA.
AC   Q4YTG6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=PB106377.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01002482; CAH98691.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   22 AA;  2696 MW;  B14A2A7BF4350AED CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 22;
  Best Local Similarity   37.5%;  Pred. No. 2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
```

```
              :::|:|:|
Db         12 TIWRILLF 19


RESULT 14
Q9R510_HELPY
ID   Q9R510_HELPY  PRELIMINARY;      PRT;    22 AA.
AC   Q9R510;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   60 kDa heat shock protein/groEL homolog (Fragment).
OS   Helicobacter pylori (Campylobacter pylori).
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Helicobacteraceae; Helicobacter.
OX   NCBI_TaxID=210;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Eschweiler B., Bohrmann B., Gerstenecker B., Schiltz E., Kist M.;
RT   "In situ localization of the 60 k protein of Helicobacter pylori,
RT   which belongs to the family of heat shock proteins, by immuno-electron
RT   microscopy.";
RL   Submitted (SEP-1994) to the EMBL/GenBank/DDBJ databases.
SQ   SEQUENCE   22 AA;  2463 MW;  117825C7D826E77B CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 22;
  Best Local Similarity   37.5%;  Pred. No. 2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          8 DSARNLLF 15


RESULT 15
Q9K8Q3_BACHD
ID   Q9K8Q3_BACHD PRELIMINARY;      PRT;    22 AA.
AC   Q9K8Q3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   BH2950 protein.
GN   OrderedLocusNames=BH2950;
OS   Bacillus halodurans.
OC   Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX   NCBI_TaxID=86665;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=C-125 / JCM 9153;
RX   MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA   Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA   Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA   Horikoshi K.;
RT   "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT   halodurans and genomic sequence comparison with Bacillus subtilis.";
RL   Nucleic Acids Res. 28:4317-4331(2000).
DR   EMBL; BA000004; BAB06669.1; -; Genomic_DNA.
DR   PIR; F84018; F84018.
KW   Complete proteome.
SQ   SEQUENCE   22 AA;  2543 MW;  12FC8AD396A81FD8 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 22;
  Best Local Similarity   37.5%;  Pred. No. 2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MKHRVLLF 8


RESULT 16
Q5BR34_SCHJA
ID   Q5BR34_SCHJA PRELIMINARY;      PRT;    23 AA.
AC   Q5BR34;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (JAN-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY915781; AAX31002.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   23 AA;  2745 MW;  4C943419F36EE4B4 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 23;
  Best Local Similarity   37.5%;  Pred. No. 2.1e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          9 WDKRSLRF 16


RESULT 17
Q4X467_PLACH
ID   Q4X467_PLACH PRELIMINARY;      PRT;    24 AA.
AC   Q4X467;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC101306.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01010161; CAH88467.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  3294 MW;  5A0496A174668617 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 24;
  Best Local Similarity   37.5%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 RNIRYLQF 18


RESULT 18
Q7M132_TREHY
ID   Q7M132_TREHY PRELIMINARY;      PRT;    24 AA.
AC   Q7M132;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
```

```
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Flagellar core protein, 34K (Fragment).
OS   Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC   Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
OX   NCBI_TaxID=159;
RN   [1]
RP   PROTEIN SEQUENCE.
RA   Koopman M.B., Baats E., van Vorstenbosch C.J., van der Zeijst B.A.,
RA   Kusters J.G.;
RT   "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT   composed of two sheath proteins and three core proteins.";
RL   J. Gen. Microbiol. 138:2697-2706(1992).
DR   PIR; C47689; C47689.
FT   NON_TER      24     24
SQ   SEQUENCE   24 AA;  2815 MW;  5EA9593CC108A0EC CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 24;
  Best Local Similarity   37.5%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NAQRTLKF 18


RESULT 19
Q4TJY7_9SPHN
ID   Q4TJY7_9SPHN PRELIMINARY;      PRT;    24 AA.
AC   Q4TJY7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=ELI2840;
OS   Erythrobacter litoralis HTCC2594.
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Sphingomonadales;
OC   Sphingomonadaceae; Erythrobacter.
OX   NCBI_TaxID=314225;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=HTCC2594;
RA   Giovannoni S.J., Cho J.-C., Ferriera S., Johnson J., Kravitz S.,
RA   Halpern A., Remington K., Beeson K., Tran B., Rogers Y.-H.,
RA   Friedman R., Venter J.C.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAGG01000008; EAL74513.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   24 AA;  2633 MW;  2CB0B0815F39C741 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 24;
  Best Local Similarity   37.5%;  Pred. No. 2.2e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 AKQRILTF 18


RESULT 20
FLAB1_TREHY
ID   FLAB1_TREHY    STANDARD;      PRT;    25 AA.
AC   P80158;
DT   01-FEB-1995 (Rel. 31, Created)
DT   01-FEB-1995 (Rel. 31, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Flagellar filament core protein flaB1 (37 kDa core protein)
DE   (Fragment).
GN   Name=flaB1;
OS   Treponema hyodysenteriae (Serpulina hyodysenteriae).
OC   Bacteria; Spirochaetes; Spirochaetales; Brachyspiraceae; Brachyspira.
OX   NCBI_TaxID=159;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=C5;
RX   MEDLINE=93139764; PubMed=1487733;
RA   Koopman M.B.H., Baats E., van Vorstenbosch C.J.A.H.V.,
RA   van der Zeijst B.A.M., Kusters J.G.;
RT   "The periplasmic flagella of Serpulina (Treponema) hyodysenteriae are
RT   composed of two sheath proteins and three core proteins.";
RL   J. Gen. Microbiol. 138:2697-2706(1992).
CC   -!- FUNCTION: Component of the core of the flagella.
CC   -!- SUBUNIT: The flagellum consists of an outer layer composed of two
CC       sheath proteins, flaA1 (44 kDa) and flaA2 (35 kDa) around a core
CC       that contains three proteins flaB1 (37 kDa), flaB2 (34 kDa) and
CC       flaB3 (32 kDa).
CC   -!- SUBCELLULAR LOCATION: Periplasmic flagellum.
CC   -!- SIMILARITY: Belongs to the bacterial flagellin family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; B47689; B47689.
KW   Direct protein sequencing; Flagellum; Periplasmic.
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2915 MW;  BBB699593CD398B6 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NAQRTLKF 18


RESULT 21
Q5C711_SCHJA
ID   Q5C711_SCHJA PRELIMINARY;      PRT;    25 AA.
AC   Q5C711;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY808674; AAX24563.1; -; mRNA.
KW   Hypothetical protein.
SQ   SEQUENCE   25 AA;  3044 MW;  E2906F3073BD1D38 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         15 IRERKLTF 22


RESULT 22
Q4X8A2_PLACH
ID   Q4X8A2_PLACH PRELIMINARY;      PRT;    25 AA.
AC   Q4X8A2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
```

```
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC405077.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01008952; CAH86874.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   25 AA;  3343 MW;  2BA053129A536078 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          6 YINRNLFF 13


RESULT 23
O11472_9HEPC
ID   O11472_9HEPC PRELIMINARY;      PRT;    25 AA.
AC   O11472;
DT   01-JUL-1997 (TrEMBLrel. 04, Created)
DT   01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   HVR1-19-1 (Fragment).
OS   Hepatitis C virus.
OC   Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC   Hepacivirus.
OX   NCBI_TaxID=11103;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=97193735; PubMed=9041320;
RA   Yoshioka K., Aiyama T., Okumura A., Takayanagi M., Iwata K.,
RA   Ishikawa T., Nagai Y., Kakumu S.;
RT   "Humoral immune response to the hypervariable region of hepatitis C
RT   virus differs between genotypes 1b and 2a.";
RL   J. Infect. Dis. 175:505-510(1997).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20013225; PubMed=10544140; DOI=10.1006/viro.1999.0004;
RA   Watanabe K., Yoshioka K., Ito H., Ishigami M., Takagi K.,
RA   Utsunomiya S., Kobayashi M., Kishimoto H., Yano M., Kakumu S.;
RT   "The hypervariable region 1 protein of hepatitis C virus broadly
RT   reactive with sera of patients with chronic hepatitis C has a similar
RT   amino acid sequence with the consensus sequence.";
RL   Virology 264:153-158(1999).
DR   EMBL; AB003918; BAA20149.1; -; Genomic_RNA.
FT   NON_TER       1      1
FT   NON_TER      25     25
SQ   SEQUENCE   25 AA;  2807 MW;  1D686B35930853C1 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          9 SRARRLSF 16


RESULT 24
Q9PRS4_ONCMY
ID   Q9PRS4_ONCMY  PRELIMINARY;      PRT;    25 AA.
AC   Q9PRS4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   C-polysaccharide binding protein 2 (Fragment).
OS   Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC   Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX   NCBI_TaxID=8022;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=96031663; PubMed=7548392;
RA   Murata M., Kodama H., Onuma M.;
RT   "Characterization of rainbow trout C-polysaccharide binding
RT   proteins.";
RL   J. Vet. Med. Sci. 57:419-425(1995).
SQ   SEQUENCE   25 AA;  3022 MW;  246C5348B32644AE CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 25;
  Best Local Similarity   37.5%;  Pred. No. 2.3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          3 PMRRSLVF 10


RESULT 25
Q4XIF3_PLACH
ID   Q4XIF3_PLACH PRELIMINARY;      PRT;    26 AA.
AC   Q4XIF3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC401025.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01005809; CAH83312.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   26 AA;  3358 MW;  E08166765C429C85 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 2.4e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          6 IKNRLLAF 13
```

```
RESULT 26
Q4XSJ0_PLACH
ID   Q4XSJ0_PLACH PRELIMINARY;      PRT;    26 AA.
AC   Q4XSJ0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC106882.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01003506; CAH80122.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   26 AA;  3306 MW;  74A4D26932F42136 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 2.4e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 LKKRSLNF 9


RESULT 27
Q91ZN8_MOUSE
ID   Q91ZN8_MOUSE PRELIMINARY;      PRT;    26 AA.
AC   Q91ZN8;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Wingless-related MMTV integration site 4 (Fragment).
GN   Name=Wnt4;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15312687; DOI=10.1016/j.canlet.2004.02.024;
RA   Peltoketo H., Allinen M., Vuosku J., Kujala S., Lundan T.,
RA   Salminen A., Winqvist R., Vainio S.;
RT   "Characterization and expression of the human WNT4; lack of associated
RT   germline mutations in high-to moderate-risk breast and ovarian
RT   cancer.";
RL   Cancer Lett. 213:83-90(2004).
DR   EMBL; AF414100; AAL10394.1; -; Genomic_DNA.
FT   NON_TER      26     26
SQ   SEQUENCE   26 AA;  2896 MW;  AD6033A9DF9501D8 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 26;
  Best Local Similarity   37.5%;  Pred. No. 2.4e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          8 RSLRLLVF 15


RESULT 28
Q7ZW70_BRARE
ID   Q7ZW70_BRARE PRELIMINARY;      PRT;    27 AA.
AC   Q7ZW70;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Zgc:77366 protein (Fragment).
GN   ORFNames=zgc:77366;
OS   Brachydanio rerio (Zebrafish) (Danio rerio).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC   Cyprinidae; Danio.
OX   NCBI_TaxID=7955;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Whole body;
RA   Strausberg R.;
RL   Submitted (APR-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC050173; AAH50173.1; -; mRNA.
DR   Ensembl; ENSDARG00000031085; Danio rerio.
DR   ZFIN; ZDB-GENE-030131-2249; zgc:77366.
FT   NON_TER       1      1
SQ   SEQUENCE   27 AA;  3373 MW;  159C3C0647ABF8F3 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity   37.5%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         10 PHRRHLYF 17


RESULT 29
Q4SKW5_TETNG
ID   Q4SKW5_TETNG PRELIMINARY;      PRT;    27 AA.
AC   Q4SKW5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF14564, whole genome shotgun sequence.
GN   ORFNames=GSTENG00016524001;
OS   Tetraodon nigroviridis (Green puffer).
```

```
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01014564; CAF98717.1; -; Genomic_DNA.
SQ   SEQUENCE   27 AA;  3278 MW;  3D551AF710009B46 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity   37.5%;  Pred. No. 2.5e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         13 TCRRPLIF 20


RESULT 30
Q7M2D7_TRYCO
ID   Q7M2D7_TRYCO PRELIMINARY;      PRT;    28 AA.
AC   Q7M2D7;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Ubiquinol-cytochrome-c reductase (EC 1.10.2.2) cytochrome b
DE   (Fragment).
OS   Trypanosoma congolense.
OG   Mitochondrion.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma;
OC   Nannomonas.
OX   NCBI_TaxID=5692;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=93382785; PubMed=8396763;
RA   Read L.K., Fish W.R., Muthiani A.M., Stuart K.;
RT   "Maxicircle DNA and edited mRNA sequences of closely related
RT   trypanosome species: implications of kRNA editing for evolution of
RT   maxicircle genomes.";
RL   Nucleic Acids Res. 21:4073-4078(1993).
DR   PIR; S41774; S41774.
DR   GO; GO:0008121; F:ubiquinol-cytochrome-c reductase activity; IEA.
FT   NON_TER      28     28
SQ   SEQUENCE   28 AA;  3493 MW;  AB4279E055308C55 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 FRCRPLLF 9


RESULT 31
Q4YJN9_PLABE
ID   Q4YJN9_PLABE PRELIMINARY;      PRT;    28 AA.
AC   Q4YJN9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PB401981.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01004442; CAI01774.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      1      1
SQ   SEQUENCE   28 AA;  3589 MW;  CEDC18ECF2689FEE CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          7 RKNRLLYF 14


RESULT 32
Q4L621_STAHJ
ID   Q4L621_STAHJ PRELIMINARY;      PRT;    28 AA.
AC   Q4L621;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Similarity.
GN   ORFNames=SH1595;
OS   Staphylococcus haemolyticus (strain JCSC1435).
OC   Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX   NCBI_TaxID=279808;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=JCSC1435;
RA   Takeuchi F., Watanabe S., Baba T., Yuzawa H., Ito T., Cui L.,
RA   Morimoto Y., Kuroda M., Takahashi M., Ankai A., Baba S., Fukui S.,
RA   Lee J.C., Hiramatsu K.;
RT   "Whole genome sequencing of Staphylococcus haemolyticus uncovers
RT   extreme plasticity of its genome and dynamism in the evolution of
RT   human-colonizing staphylococcal species.";
RL   Submitted (DEC-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP006716; BAE04904.1; -; Genomic_DNA.
SQ   SEQUENCE   28 AA;  3432 MW;  0777E6A2E541B173 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 XXXRXLXF 8
              :::|:|:|
Db          4 ELRRKLTF 11


RESULT 33
Q9DD70_CHICK
ID   Q9DD70_CHICK PRELIMINARY;      PRT;    28 AA.
AC   Q9DD70;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Growth hormone receptor (Fragment).
GN   Name=GHR;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Li L., Wang X., Cogburn L.A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254949; AAG10253.1; -; mRNA.
DR   EMBL; AF254951; AAG10251.1; -; Genomic_DNA.
DR   EMBL; AF254952; AAG10252.1; -; Genomic_DNA.
DR   GO; GO:0004872; F:receptor activity; IEA.
KW   Receptor.
FT   NON_TER      28     28
SQ   SEQUENCE   28 AA;  3162 MW;  507F235BC629D722 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MDLRHLLF 8


RESULT 34
Q9DFT7_CHICK
ID   Q9DFT7_CHICK PRELIMINARY;      PRT;    28 AA.
AC   Q9DFT7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Growth hormone receptor (Fragment).
GN   Name=GHR;
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Liver;
RA   Li L., Wang X., Cogburn L.A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF254950; AAG10250.1; -; Genomic_DNA.
DR   GO; GO:0004872; F:receptor activity; IEA.
KW   Receptor.
FT   NON_TER      28     28
SQ   SEQUENCE   28 AA;  3146 MW;  447E635BC629D722 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 28;
  Best Local Similarity   37.5%;  Pred. No. 2.6e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 MDLRHLLF 8


RESULT 35
AP65_CARMA
ID   AP65_CARMA     STANDARD;      PRT;    30 AA.
AC   P82964;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Antibacterial 6.5 kDa protein (Fragment).
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   PROTEIN SEQUENCE, AND FUNCTION.
RC   TISSUE=Hemocyte;
RX   MEDLINE=97008941; PubMed=8856051;
RA   Schnapp D., Kemp G.D., Smith V.J.;
RT   "Purification and characterization of a proline-rich antibacterial
RT   peptide, with sequence similarity to bactenecin-7, from the haemocytes
RT   of the shore crab, Carcinus maenas.";
RL   Eur. J. Biochem. 240:532-539(1996).
CC   -!- FUNCTION: Strong antimicrobial activity against P.immobilis and
CC       M.luteus, less active against E.coli D22.
CC   -!- MISCELLANEOUS: On the 2D-gel the determined MW is: 6.5 kDa.
CC   -!- SIMILARITY: To bovine bactenecin 7.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; S74112; S74112.
KW   Antibiotic; Antimicrobial; Direct protein sequencing.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3307 MW;  6E2C2205934896C4 CRC64;

  Query Match             100.0%;  Score 20;  DB 1;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         14 IGPRPLPF 21


RESULT 36
Q4XCS4_PLACH
ID   Q4XCS4_PLACH PRELIMINARY;      PRT;    30 AA.
AC   Q4XCS4;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC403287.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
```

```
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01007501; CAH85298.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   30 AA;  3782 MW;  CB4EC601B0BDF7E0 CRC64;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         13 LFCRLLIF 20


RESULT 37
Q4YQ95_PLABE
ID   Q4YQ95_PLABE PRELIMINARY;      PRT;    30 AA.
AC   Q4YQ95;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PB107776.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01003005; CAH99814.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   30 AA;  3718 MW;  A608D7A02BCC490F CRC64;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          5 YTSRALSF 12


RESULT 38
Q9R5K3_LEPIN
ID   Q9R5K3_LEPIN  PRELIMINARY;      PRT;    30 AA.
AC   Q9R5K3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   34 kDa periplasmic flagella core protein (Fragment).
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92325069; PubMed=1624463;
RA   Trueba G.A., Bolin C.A., Zuerner R.L.;
RT   "Characterization of the periplasmic flagellum proteins of Leptospira
RT   interrogans.";
RL   J. Bacteriol. 174:4761-4768(1992).
DR   PIR; A44913; A44913.
SQ   SEQUENCE   30 AA;  3464 MW;  3BDDEF3FD5CA4969 CRC64;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 NSHRVLKF 18


RESULT 39
Q9RER6_ENTAE
ID   Q9RER6_ENTAE PRELIMINARY;      PRT;    30 AA.
AC   Q9RER6;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Putative AmpR protein (Fragment).
GN   Name=ampR;
OS   Enterobacter aerogenes (Aerobacter aerogenes).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Enterobacter.
OX   NCBI_TaxID=548;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=20493130; PubMed=11036041;
RX   DOI=10.1128/AAC.44.11.3158-3162.2000;
RA   Preston K.E., Radomski C.C.A., Venezia R.A.;
RT   "Nucleotide sequence of the chromosomal ampC gene of Enterobacter
RT   aerogenes.";
RL   Antimicrob. Agents Chemother. 44:3158-3162(2000).
DR   EMBL; AF211348; AAF18993.1; -; Genomic_DNA.
DR   GO; GO:0003700; F:transcription factor activity; IEA.
DR   GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR   InterPro; IPR000847; HTH_LysR.
DR   PROSITE; PS50931; HTH_LYSR; 1.
FT   NON_TER      30     30
SQ   SEQUENCE   30 AA;  3407 MW;  D6A2370BE3D5B7C5 CRC64;

  Query Match           100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         16 AAARHLSF 23


RESULT 40
Q57H84_SALCH
ID   Q57H84_SALCH PRELIMINARY;      PRT;    30 AA.
AC   Q57H84;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SC3333, SC4022, SC4057;
OS   Salmonella cholerae-suis (Salmonella enterica).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=591;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=SC-B67;
RX   PubMed=15781495;
RA   Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA   Wang H.-S., Lee Y.-S.;
RT   "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT   highly invasive and resistant zoonotic pathogen.";
RL   Nucleic Acids Res. 33:1690-1698(2005).
```

```
DR   EMBL; AE017220; AAX67928.1; -; Genomic_DNA.
DR   EMBL; AE017220; AAX67963.1; -; Genomic_DNA.
DR   EMBL; AE017220; AAX67239.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   30 AA;  3502 MW;  42E7160A3DEA5F27 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         15 SRLRYLLF 22


RESULT 41
Q5PFE0_SALPA
ID   Q5PFE0_SALPA PRELIMINARY;      PRT;    30 AA.
AC   Q5PFE0;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SPA2518;
OS   Salmonella paratyphi-a.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=54388;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 9150;
RX   PubMed=15531882; DOI=10.1038/ng1470;
RA   McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA   Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA   Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA   Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA   Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA   Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA   Spieth J., Wilson R.K.;
RT   "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT   restricted serovars of Salmonella enterica that cause typhoid.";
RL   Nat. Genet. 36:1268-1274(2004).
DR   EMBL; CP000026; AAV78389.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   30 AA;  3639 MW;  ECD8CB180EEB5F32 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 30;
  Best Local Similarity   37.5%;  Pred. No. 2.8e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         15 SRLRYLLF 22


RESULT 42
Q4XBW6_PLACH
ID   Q4XBW6_PLACH PRELIMINARY;      PRT;    31 AA.
AC   Q4XBW6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC403662.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01007789; CAH85606.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   31 AA;  4064 MW;  7A2CD8A06505F5BC CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          9 FIIRKLFF 16


RESULT 43
Q4KA34_PSEF5
ID   Q4KA34_PSEF5 PRELIMINARY;      PRT;    31 AA.
AC   Q4KA34;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=PFL_3799;
OS   Pseudomonas fluorescens (strain Pf-5).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=220664;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Pf-5;
RX   PubMed=15980861; DOI=10.1038/nbt1110;
RA   Paulsen I.T., Press C., Ravel J., Kobayashi D., Myers G.S.,
RA   Mavrodi D., DeBoy R.T., Seshadri R., Ren Q., Madupu R., Dodson R.J.,
RA   Durkin S., Brinkac L.M., Daugherty S.C., Sullivan S.A., Rosovitz M.,
RA   Gwinn M.L., Zhou L., Nelson W.C., Weidman J., Watkins K., Tran K.,
RA   Khouri H.M., Pierson E., Pierson L. III, Thomashow L., Loper J.;
RT   "Complete genome sequence of the plant commensal Pseudomonas
RT   fluorescens Pf-5.";
RL   Nat. Biotechnol. 23:873-878(2005).
DR   EMBL; CP000076; AAY93062.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   31 AA;  3593 MW;  45D45B4E6AE51501 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         21 LGDRHLIF 28


RESULT 44
Q65TM8_MANSM
ID   Q65TM8_MANSM PRELIMINARY;      PRT;    31 AA.
AC   Q65TM8;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MS1075;
OS   Mannheimia succiniciproducens (strain MBEL55E).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;
OC   Pasteurellaceae; Mannheimia.
OX   NCBI_TaxID=221988;
```

```
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15378067; DOI=10.1038/nbt1010;
RA   Hong S.H., Kim J.S., Lee S.Y., In Y.H., Choi S.S., Rih J.-K.,
RA   Kim C.H., Jeong H., Hur C.G., Kim J.J.;
RT   "The genome sequence of the capnophilic rumen bacterium Mannheimia
RT   succiniciproducens.";
RL   Nat. Biotechnol. 22:1275-1281(2004).
DR   EMBL; AE016827; AAU37682.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   31 AA;  4164 MW;  A55BB45BCDF46339 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         12 RLERRLIF 19


RESULT 45
Q87K91_VIBPA
ID   Q87K91_VIBPA PRELIMINARY;      PRT;    31 AA.
AC   Q87K91;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein VPA0007.
GN   OrderedLocusNames=VPA0007;
OS   Vibrio parahaemolyticus.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC   Vibrionaceae; Vibrio.
OX   NCBI_TaxID=670;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=RIMD 2210633 / Serotype O3:K6;
RX   MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA   Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA   Iijima Y., Najima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA   Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT   "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT   distinct from that of V. cholerae.";
RL   Lancet 361:743-749(2003).
DR   EMBL; BA000032; BAC61350.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   31 AA;  3814 MW;  2A278800BF7D7C2A CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         24 ASLRRLAF 31


RESULT 46
Q8KBJ8_CHLTE
ID   Q8KBJ8_CHLTE PRELIMINARY;      PRT;    31 AA.
AC   Q8KBJ8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=CT1789;
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobaculum.
OX   NCBI_TaxID=1097;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA   Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA   Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA   Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA   Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR   EMBL; AE006470; AAM73010.1; -; Genomic_DNA.
DR   TIGR; CT1789; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   31 AA;  3827 MW;  4EF41565B6B535BE CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         10 ELFRKLFF 17


RESULT 47
Q67974_HPBV0
ID   Q67974_HPBV0 PRELIMINARY;      PRT;    31 AA.
AC   Q67974;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   X protein (Fragment).
OS   Hepatitis B virus.
OC   Viruses; Retroid viruses; Hepadnaviridae; Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85270; CAA59557.1; -; Genomic_DNA.
DR   PIR; S53153; S53153.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR000236; TransactX.
DR   Pfam; PF00739; X; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   31 AA;  3288 MW;  F1E9DC3EF0261B20 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 ADSRLLIF 9


RESULT 48
Q68005_HPBV0
ID   Q68005_HPBV0 PRELIMINARY;      PRT;    31 AA.
AC   Q68005;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   X protein (Fragment).
GN   Name=X;
OS   Hepatitis B virus.
OC   Viruses; Retro-transcribing viruses; Hepadnaviridae;
OC   Orthohepadnavirus.
OX   NCBI_TaxID=10407;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Lai M.E., Mazzoleni A.P., Porru A., Balestrieri A.;
```

```
RL   Submitted (MAR-1995) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X85257; CAA59520.1; -; Genomic_DNA.
DR   PIR; S53192; S53192.
DR   GO; GO:0019079; P:viral genome replication; IEA.
DR   InterPro; IPR000236; TransactX.
DR   Pfam; PF00739; X; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   31 AA;  3377 MW;  B28DDC2000D8EA72 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 31;
  Best Local Similarity   37.5%;  Pred. No. 2.9e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          2 EELRLLVF 9


RESULT 49
Q9BYF3_HUMAN
ID   Q9BYF3_HUMAN PRELIMINARY;      PRT;    32 AA.
AC   Q9BYF3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Ribosomal protein L36 (Fragment).
GN   Name=RPL36;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=21295043; PubMed=11401437; DOI=10.1006/geno.2000.6470;
RA   Uechi T., Tanaka T., Kenmochi N.;
RT   "A complete map of the human ribosomal protein genes: assignment of 80
RT   genes to the cytogenetic map and implications for human disorders.";
RL   Genomics 72:223-230(2001).
DR   EMBL; AB046410; BAB21256.1; -; Genomic_DNA.
DR   GO; GO:0005840; C:ribosome; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   GO; GO:0006412; P:protein biosynthesis; IEA.
DR   InterPro; IPR000509; Ribosomal_L36e.
DR   PANTHER; PTHR10114; Ribosomal_L36e; 1.
DR   Pfam; PF01158; Ribosomal_L36e; 1.
DR   ProDom; PD009192; Ribosomal_L36e; 1.
KW   Ribonucleoprotein; Ribosomal protein.
FT   NON_TER       1       1
FT   NON_TER      32      32
SQ   SEQUENCE   32 AA;  3835 MW;  C55DC318353C2C3A CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 32;
  Best Local Similarity   37.5%;  Pred. No. 3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          5 KDKRALKF 12


RESULT 50
Q4X2P5_PLACH
ID   Q4X2P5_PLACH PRELIMINARY;      PRT;    32 AA.
AC   Q4X2P5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PC405872.00.0;
OS   Plasmodium chabaudi.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5825;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAJ01010606; CAH89088.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1       1
SQ   SEQUENCE   32 AA;  3788 MW;  8971907C3EC3B3B0 CRC64;

  Query Match             100.0%;  Score 20;  DB 2;  Length 32;
  Best Local Similarity   37.5%;  Pred. No. 3e+03;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db         11 VYVRLLPF 18


Search completed: May  5, 2006, 12:23:33
Job time : 80 secs
```

GenCore version 5.1.7

OM protein - protein search, using sw model

Run on: May 5, 2006, 12:18:41 ; Search time 187 Seconds
(without alignments)
18.797 Million cell updates/sec

Title: US-09-726-470A-2
Perfect score: 20
Sequence: 1 XXXRXLXF 8

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 100 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	100.0	8	2	AAR89217	Aar89217 CSF clone
2	20	100.0	8	2	AAW57008	Aaw57008 Enzyme in
3	20	100.0	8	2	AAW64283	Aaw64283 mMCP-7 pe
4	20	100.0	8	4	AAG66260	Aag66260 p21 C-ter
5	20	100.0	8	4	AAG66252	Aag66252 p21 C-ter
6	20	100.0	8	4	AAG66271	Aag66271 p21 C-ter
7	20	100.0	8	4	AAU05707	Aau05707 p21 C-ter
8	20	100.0	8	4	AAU05736	Aau05736 p21 C-ter
9	20	100.0	8	4	AAG66202	Aag66202 p21 deriv
10	20	100.0	8	4	AAG66256	Aag66256 p21 C-ter
11	20	100.0	8	4	AAG66258	Aag66258 p21 C-ter
12	20	100.0	8	4	AAG66275	Aag66275 p21 C-ter
13	20	100.0	8	4	AAU05708	Aau05708 p21 C-ter
14	20	100.0	8	4	AAG66201	Aag66201 p21 deriv
15	20	100.0	8	4	AAG66274	Aag66274 p21 C-ter
16	20	100.0	8	4	AAU05710	Aau05710 p21 C-ter
17	20	100.0	8	4	AAU05742	Aau05742 p21 C-ter
18	20	100.0	8	4	AAG66257	Aag66257 p21 C-ter
19	20	100.0	8	4	AAG66259	Aag66259 p21 C-ter
20	20	100.0	8	4	AAG66261	Aag66261 p21 C-ter
21	20	100.0	8	4	AAG66273	Aag66273 p21 C-ter
22	20	100.0	8	4	AAG65137	Aag65137 Synthetic
23	20	100.0	8	4	AAG66262	Aag66262 p21 C-ter
24	20	100.0	8	4	AAG66264	Aag66264 p21 C-ter
25	20	100.0	8	4	AAG66265	Aag66265 p21 C-ter
26	20	100.0	8	4	AAG65145	Aag65145 Synthetic
27	20	100.0	8	4	AAG66207	Aag66207 p21 deriv
28	20	100.0	8	4	AAG66263	Aag66263 p21 C-ter
29	20	100.0	8	4	AAG65127	Aag65127 p21WAFI C
30	20	100.0	8	4	AAU05706	Aau05706 p21 C-ter
31	20	100.0	8	4	AAU05741	Aau05741 p21 C-ter
32	20	100.0	8	4	AAG66253	Aag66253 p21 C-ter
33	20	100.0	8	4	AAG66272	Aag66272 p21 C-ter
34	20	100.0	8	4	AAG66203	Aag66203 p21 deriv
35	20	100.0	8	4	AAU05740	Aau05740 p21 C-ter
36	20	100.0	8	4	AAG66205	Aag66205 p21 deriv
37	20	100.0	8	4	AAU05709	Aau05709 p21 C-ter
38	20	100.0	8	4	AAG65150	Aag65150 p21 deriv
39	20	100.0	8	4	AAG66266	Aag66266 p21 C-ter
40	20	100.0	8	4	AAJ02281	Aaj02281 Hepatitis
41	20	100.0	8	8	ADJ72146	Adj72146 Cyclin re
42	20	100.0	8	8	ADJ72150	Adj72150 Cyclin re
43	20	100.0	8	9	ADZ77243	Adz77243 Rapamycin
44	20	100.0	8	9	ADZ71454	Adz71454 p21-deriv
45	20	100.0	8	9	ADZ71458	Adz71458 p21-deriv
46	20	100.0	8	9	ADZ71490	Adz71490 p21-deriv
47	20	100.0	8	9	ADZ71491	Adz71491 p21-deriv
48	20	100.0	8	9	ADZ71494	Adz71494 p21-deriv
49	20	100.0	8	9	ADZ71504	Adz71504 p21-deriv
50	20	100.0	8	9	ADZ71838	Adz71838 p21-deriv
51	20	100.0	8	9	ADZ71854	Adz71854 p21-deriv
52	20	100.0	8	9	ADZ71979	Adz71979 p21-deriv
53	20	100.0	8	9	ADZ71447	Adz71447 p21-deriv
54	20	100.0	8	9	ADZ71573	Adz71573 p21-deriv
55	20	100.0	8	9	ADZ71593	Adz71593 p21-deriv
56	20	100.0	8	9	ADZ71596	Adz71596 p21-deriv
57	20	100.0	8	9	ADZ71870	Adz71870 p21-deriv
58	20	100.0	8	9	ADZ71440	Adz71440 p21-deriv
59	20	100.0	8	9	ADZ71471	Adz71471 p21-deriv
60	20	100.0	8	9	ADZ71487	Adz71487 p21-deriv
61	20	100.0	8	9	ADZ71572	Adz71572 p21-deriv
62	20	100.0	8	9	ADZ71598	Adz71598 p21-deriv
63	20	100.0	8	9	ADZ71778	Adz71778 p21-deriv
64	20	100.0	8	9	ADZ71496	Adz71496 p21-deriv
65	20	100.0	8	9	ADZ71509	Adz71509 p21-deriv
66	20	100.0	8	9	ADZ71570	Adz71570 p21-deriv
67	20	100.0	8	9	ADZ71597	Adz71597 p21-deriv
68	20	100.0	8	9	ADZ71855	Adz71855 p21-deriv
69	20	100.0	8	9	ADZ71868	Adz71868 p21-deriv
70	20	100.0	8	9	ADZ71879	Adz71879 p21-deriv
71	20	100.0	8	9	ADZ71899	Adz71899 p21-deriv
72	20	100.0	8	9	ADZ71978	Adz71978 p21-deriv
73	20	100.0	8	9	ADZ71478	Adz71478 p21-deriv
74	20	100.0	8	9	ADZ71505	Adz71505 p21-deriv
75	20	100.0	8	9	ADZ71774	Adz71774 p21-deriv
76	20	100.0	8	9	ADZ71856	Adz71856 p21-deriv
77	20	100.0	8	9	ADZ71863	Adz71863 p21-deriv
78	20	100.0	8	9	ADZ71865	Adz71865 p21-deriv
79	20	100.0	8	9	ADZ71976	Adz71976 p21-deriv
80	20	100.0	8	9	ADZ71466	Adz71466 p21-deriv
81	20	100.0	8	9	ADZ71851	Adz71851 p21-deriv
82	20	100.0	8	9	ADZ71885	Adz71885 p21-deriv
83	20	100.0	8	9	ADZ71973	Adz71973 p21-deriv
84	20	100.0	8	9	ADZ71977	Adz71977 p21-deriv
85	20	100.0	8	9	ADZ71980	Adz71980 p21-deriv
86	20	100.0	8	9	ADZ71981	Adz71981 p21-deriv
87	20	100.0	8	9	ADZ71453	Adz71453 p21-deriv
88	20	100.0	8	9	ADZ71464	Adz71464 p21-deriv
89	20	100.0	8	9	ADZ71468	Adz71468 p21-deriv
90	20	100.0	8	9	ADZ71492	Adz71492 p21-deriv
91	20	100.0	8	9	ADZ71412	Adz71412 p21-deriv
92	20	100.0	8	9	ADZ71461	Adz71461 p21-deriv
93	20	100.0	8	9	ADZ71467	Adz71467 p21-deriv
94	20	100.0	8	9	ADZ71476	Adz71476 p21-deriv
95	20	100.0	8	9	ADZ71599	Adz71599 p21-deriv
96	20	100.0	8	9	ADZ71847	Adz71847 p21-deriv
97	20	100.0	8	9	ADZ71864	Adz71864 p21-deriv

98	20	100.0	8	9	ADZ71472	Adz71472 p21-deriv
99	20	100.0	8	9	ADZ71477	Adz71477 p21-deriv
100	20	100.0	8	9	ADZ71488	Adz71488 p21-deriv

ALIGNMENTS

RESULT 1
AAR89217
ID AAR89217 standard; peptide; 8 AA.
XX
AC AAR89217;
XX
DT 05-SEP-1996 (first entry)
XX
DE CSF clone A10 Vbeta8-CDR3.
XX
KW Polymerase chain reaction; PCR; primer; amplify; human; T cell receptor;
KW beta chain; TCR; myelin basic protein; BP; autoantigen; encephalitogen;
KW experimental autoimmune encephalomyelitis; EAE; multiple sclerosis; MS;
KW autoimmune disease; neurological disease; cerebrospinal fluid; therapy;
KW central nervous system; complementarity determining region; CDR;
KW T lymphocyte; optical nerve damage; anterior chamber inflammation.
XX
OS Synthetic.
XX
PN WO9601329-A1.
XX
PD 18-JAN-1996.
XX
PF 26-JUN-1995; 95WO-US008086.
XX
PR 01-JUL-1994; 94US-00270634.
XX
PA (CONN-) CONNECTIVE THERAPEUTICS INC.
XX
PI Vandenbark AA, Offner H, Buenafe A;
XX
DR WPI; 1996-087679/09.
DR N-PSDB; AAT10592.
XX
PT Methods for diagnosis and immune-related therapy of autoimmune diseases -
PT partic. multiple sclerosis, by detecting marker T cell receptor V gene
PT bias and treating patients with selected V beta peptide(s).
XX
PS Example 2; Fig 4a; 62pp; English.
XX
CC AAR89215-R89251 represent clones of the Vbeta8 complementarity
CC determining region 3 (CDR3) of the T cell receptor beta (TCRbeta) chain.
CC These sequences were isolated from cerebrospinal fluid (CSF), spinal cord
CC (SC) and lymph nodes (LN) of clones of Lewis rats with experimental
CC autoimmune encephalomyelitis (EAE). By detecting the presence of a marker
CC TCR V gene bias in a body fluid which encapsulates all or part of the
CC target organ, an autoimmune disease (such as a neurological disease) in a
CC human can be identified. This method can also be carried out to detect
CC the presence of a biased motif common to T cell receptors specific for
CC the pathogenic antigen in a non-target tissue or organ. By analysing the
CC Vbeta gene repertoire of CSF, and determining the presence of a Vbeta
CC gene bias, an immune-related disease that targets the central nervous
CC system can be diagnosed. Therapeutic Vbeta peptide sequences can be
CC selected to use as treatment of a disease or condition. The selection is
CC carried out by identifying a Vbeta gene bias in a body fluid that is not
CC the target tissue or organ of the disease, and selecting an immunogenic
CC peptide corresponding to the Vbeta gene bias. Multiple sclerosis (MS) can
CC be treated by identifying the CDR2 of a V gene peptide on the surface of
CC a T lymphocyte in the CSF of a patient and administering a peptide
CC corresponding to this region. These methods can also be used for the
CC diagnosis and immune-related therapy of optical nerve damage and anterior
CC chamber inflammation as well as other human neurological diseases
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 2; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 DSTRALYF 8

RESULT 2
AAW57008
ID AAW57008 standard; peptide; 8 AA.
XX
AC AAW57008;
XX
DT 28-JUL-1998 (first entry)
XX
DE Enzyme inhibitor peptide SEQ ID NO:209.
XX
KW Enzyme inhibitor; t-PA; u-PA; chymotrypsin; serine protease; active;
KW latent; substrate subtraction phage display peptide library;
KW identification; kinase; phosphatase; serpin.
XX
OS Homo sapiens.
XX
PN WO9747314-A1.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US009760.
XX
PR 10-JUN-1996; 96US-0019495P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Madison EL, Ke S;
XX
DR WPI; 1998-062746/06.
XX
PT Substrate subtraction phage display peptide libraries - used to
PT distinguish between active and latent forms of enzyme, e.g. serine
PT protease.
XX
PS Claim 25; Page 111; 138pp; English.
XX
CC The present sequence represents an enzyme inhibitor peptide used in the
CC method of the invention to distinguish between t-PA and u-PA. The present
CC invention describes a substrate subtraction library for the
CC identification of peptide substrates selective between a first enzyme
CC (E1) and a second enzyme (E2), comprising a collection different
CC peptides, substantially lacking peptides that are effective substrates
CC for E1. Also described are: (1) a method (M1) for identifying peptide
CC substrates selective between a first enzyme (E1) and a second enzyme (E2)
CC ; (2) a compound comprising the amino acid sequence of a peptide
CC identified by M1; (3) a polypeptide for use as an enzyme inhibitor
CC comprising one of 237 amino acid sequences (see AAW56801 to AAW56947, and
CC AAW56949 to AAW57038); (4) a recombinant DNA vector comprising DNA (I)
CC encoding a protease inhibitor including the sequence identified by the M1
CC ; (5) a prokaryotic or eukaryotic cell containing the vector of (4); (6)
CC an antibody (Ab) immunoreactive with at least one of the peptides
CC identified by M1; and (7) a diagnostic assay for distinguishing between
CC active and latent forms of protease inhibitors, that uses (Ab). The
CC library and method are used for distinguishing between active and latent
CC forms of enzyme inhibitors, e.g. proteases, kinases and phosphatases.
CC (Ab) are used for affinity purification of recombinant peptides and in
CC the identification of naturally occurring protease inhibitors. Enzyme-
CC inhibiting peptides identified can be used to treat a serpin deficiency
CC or a disorder of serine proteases
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 2; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;

```
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 SGRRTLDF 8


RESULT 3
AAW64283
ID   AAW64283 standard; peptide; 8 AA.
XX
AC   AAW64283;
XX
DT   24-NOV-1998  (first entry)
XX
DE   mMCP-7 peptide substrate.
XX
KW   MCP-7; mouse; mast cell protease 7; tryptase-7; blood clot;
KW   anticoagulant; myocardial infarction; reocclusion; thromboembolism;
KW   cerebral embolism; thrombosis; therapy.
XX
OS   Synthetic.
XX
PN   WO9824886-A1.
XX
PD   11-JUN-1998.
XX
PF   25-NOV-1997;   97WO-US021620.
XX
PR   04-DEC-1996;   96US-0032354P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL.
XX
PI   Stevens RL;
XX
DR   WPI; 1998-333308/29.
XX
PT   New compositions containing tryptase-7, e.g. mouse mast cell protease-7 -
PT   are used to treat clot formation in e.g. myocardial infarction,
PT   reocclusion following angioplasty or pulmonary thrombo-embolism.
XX
PS   Example; Page 46; 92pp; English.
XX
CC   This is a substrate peptide of mouse mast cell protease 7 (mMCP-7, see
CC   AAW64233). It is one of 21 peptides (see AAW64270-90) obtained by
CC   incubating a phage display peptide library 2 times with a recombinant
CC   FLAG-tagged mMCP-7 polypeptide, isolating clones, and deducing the amino
CC   acid sequence of the protease susceptible domains in the pIII fusion
CC   proteins. Only one peptide (see AAW64270) was obtained after 4 rounds of
CC   screening. mMCP-7 has been characterised as having fibrinogen as its
CC   physiological substrate. It can be used to prevent or treat fibrin clot
CC   formation in vitro and in vivo. Tryptase-7 proteases of the invention,
CC   including mMCP-7 and its homologues, can be used to treat disorders that
CC   are mediated by undesirable thrombus clot formation, such as myocardial
CC   infarct and reocclusion following angioplasty, and are also useful for
CC   surgical procedures that require that blood does not clot
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 2;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 LSYRKLRF 8


RESULT 4
AAG66260
ID   AAG66260 standard; peptide; 8 AA.
XX
AC   AAG66260;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #52.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HVKRRLIF 8


RESULT 5
AAG66252
```

```
ID   AAG66252 standard; peptide; 8 AA.
XX
AC   AAG66252;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #44.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 AAKRRLIF 8
```

```
RESULT 6
AAG66271
ID   AAG66271 standard; peptide; 8 AA.
XX
AC   AAG66271;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #63.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
```

```
              :::|:|:|
Db          1 HAKRALIF 8


RESULT 7
AAU05707
ID   AAU05707 standard; protein; 8 AA.
XX
AC   AAU05707;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #74.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLLF 8


RESULT 8
AAU05736
ID   AAU05736 standard; protein; 8 AA.
XX
AC   AAU05736;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #106.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   7
FT                   /label= OTHER
FT                   /note= "Other= para-fluorophenylalanine,
FT                   dichlorophenylalanine, para-chlorophenylalanine, meta-
FT                   chlorophenylalanine, ortho-chlorophenylalanine, TyrMe
FT                   (not defined), thiophenylalanine or 3-pyridylalanine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 90; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
```

```
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|||
Db          1 AAKRRLXF 8


RESULT 9
AAG66202
ID   AAG66202 standard; peptide; 8 AA.
XX
AC   AAG66202;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived peptide, p21(152)Ser153Ala #2.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21(152)Ser153Ala.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Misc-difference 2
FT                   /label= Ala, Gly, Abu, Val, Ile, Phe, Nva, OTHER
FT                   /note= "Other= phenylglycine, t-butylglycine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 14; Page 54; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21 and used in a Cyclin A binding experiment, the effect
CC   on cyclin A binding of replacing the Ala residue at position 2 was
CC   assessed
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :|:|:|:|
Db          1 HXLRRLIF 8


RESULT 10
AAG66256
ID   AAG66256 standard; protein; 8 AA.
XX
AC   AAG66256;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #49.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
```

```
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 KACRRLIF 8


RESULT 11
AAG66258
ID   AAG66258 standard; peptide; 8 AA.
XX
AC   AAG66258;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #50.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HGKRRLIF 8


RESULT 12
AAG66275
ID   AAG66275 standard; peptide; 8 AA.
XX
AC   AAG66275;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #67.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   5
FT                   /label= OTHER, Aib
FT                   /note= "Other= Citrulline or sarcosine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
```

```
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|||:|
Db          1 HAKRXLIF 8


RESULT 13
AAU05708
ID   AAU05708 standard; protein; 8 AA.
XX
AC   AAU05708;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #75.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLVF 8


RESULT 14
AAG66201
ID   AAG66201 standard; peptide; 8 AA.
XX
AC   AAG66201;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived peptide, p21(152)Ser153Ala.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21(152)Ser153Ala.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /label= Ala, Phe, OTHER
FT                   /note= "Other= 3-pyridylalanine,Thiophenylalanine,
FT                   Homoserine, 2,3-Diaminobutyric acid or absent and the N-
FT                   terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
```

DR WPI; 2001-488493/53.
XX
PT New p21 derived peptides and their variants, particularly useful as
PT selective inhibitors of CDK2/cyclin interaction for treating
PT proliferative disorders e.g. cancers and leukemias, and in assays for
PT identifying CDK/cyclin inhibitors.
XX
PS Example 13; Page 54; 102pp; English.
XX
CC The invention relates to peptide and their variants derived from p21WAFI,
CC which are inhibitors of CDK2 activity by binding to G1 and S phase
CC specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC peptide may have further amino acids at either end or have up to 7 amino
CC acids deleted, provided the motif XLXF is retained. The peptides are
CC specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC preferably cyclins which activate CDK2. One of the peptides corresponds
CC to p21(149-159). The peptides are used for treating proliferative
CC disorders, e.g. cancers and leukaemias. The peptides are also for
CC identifying substances which interfere with protein-protein interactions
CC involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC activity. P21 peptides other than p21(149-159) competitively inhibit the
CC binding of peptide p21(149-159) to cyclin and may be used to identify
CC substances that bind to, or inhibit peptide- cyclin interactions.
CC Substances for screening in the assays include antibody products specific
CC for p21 or cyclin binding regions, combinatorial libraries and single
CC compound collections. The present sequence is a peptide derived from the
CC C-terminus of p21 and used in a Cyclin A binding experiment, the effect
CC on cyclin A binding of replacing the His residue at position 1 was
CC assessed
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 4; Length 8;
Best Local Similarity 50.0%; Pred. No. 2e+06;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XXXRXLXF 8
              |::|:|:|
Db          1 XAKRRLIF 8
```

RESULT 15
AAG66274
ID AAG66274 standard; peptide; 8 AA.
XX
AC AAG66274;
XX
DT 21-NOV-2001 (first entry)
XX
DE p21 C-terminus derived peptide #66.
XX
KW Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "The N-terminus is hydrogenated"
FT Modified-site 8
FT /note= "C-terminal amide"
XX
PN WO200140142-A2.
XX
PD 07-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-GB004550.
XX
PR 30-NOV-1999; 99GB-00028323.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2001-488493/53.
XX
PT New p21 derived peptides and their variants, particularly useful as
PT selective inhibitors of CDK2/cyclin interaction for treating
PT proliferative disorders e.g. cancers and leukemias, and in assays for
PT identifying CDK/cyclin inhibitors.
XX
PS Claim 25; Page 88; 102pp; English.
XX
CC The invention relates to peptide and their variants derived from p21WAFI,
CC which are inhibitors of CDK2 activity by binding to G1 and S phase
CC specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC peptide may have further amino acids at either end or have up to 7 amino
CC acids deleted, provided the motif XLXF is retained. The peptides are
CC specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC preferably cyclins which activate CDK2. One of the peptides corresponds
CC to p21(149-159). The peptides are used for treating proliferative
CC disorders, e.g. cancers and leukaemias. The peptides are also for
CC identifying substances which interfere with protein-protein interactions
CC involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC activity. P21 peptides other than p21(149-159) competitively inhibit the
CC binding of peptide p21(149-159) to cyclin and may be used to identify
CC substances that bind to, or inhibit peptide- cyclin interactions.
CC Substances for screening in the assays include antibody products specific
CC for p21 or cyclin binding regions, combinatorial libraries and single
CC compound collections. The present sequence is a peptide derived from the
CC C-terminus of p21
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 4; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRSLIF 8
```

RESULT 16
AAU05710
ID AAU05710 standard; protein; 8 AA.
XX
AC AAU05710;
XX
DT 21-NOV-2001 (first entry)
XX
DE p21 C-terminus derived peptide #77.
XX
KW Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "The N-terminus is hydrogenated"
FT Modified-site 8
FT /note= "C-terminal amide"
XX
PN WO200140142-A2.
XX
PD 07-JUN-2001.
XX
PF 29-NOV-2000; 2000WO-GB004550.

```
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLFF 8


RESULT 17
AAU05742
ID   AAU05742 standard; protein; 8 AA.
XX
AC   AAU05742;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #111.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 34; Page 92; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLLF 8


RESULT 18
AAG66257
ID   AAG66257 standard; peptide; 8 AA.
XX
AC   AAG66257;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #49.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
```

```
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 FAKRRLIF 8


RESULT 19
AAG66259
ID   AAG66259 standard; peptide; 8 AA.
XX
AC   AAG66259;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #51.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   2
FT                   /label= OTHER, Abu, Nva
FT                   /note= "Other= t-Butylglycine or phenylglycine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :|:|:|:|
Db          1 HXKRRLIF 8


RESULT 20
AAG66261
ID   AAG66261 standard; peptide; 8 AA.
XX
AC   AAG66261;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #53.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
```

```
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HIKRRLIF 8


RESULT 21
AAG66273
ID   AAG66273 standard; peptide; 8 AA.
XX
AC   AAG66273;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #65.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRPLIF 8


RESULT 22
AAG65137
ID   AAG65137 standard; peptide; 8 AA.
XX
AC   AAG65137;
XX
```

```
DT   21-NOV-2001  (first entry)
XX
DE   Synthetic peptide, p21 C-terminus (S153A).
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21 C-terminus (S153A).
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "Optional Hydrogenated N-terminus"
FT   Modified-site   8
FT                   /note= "Optional C-terminal carboxamide or amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 15; Page 84; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a synthetic peptide derived
CC   from the C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLIF 8


RESULT 23
AAG66262
```

```
ID   AAG66262 standard; peptide; 8 AA.
XX
AC   AAG66262;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #54.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HFKRRLIF 8
```

```
RESULT 24
AAG66264
ID   AAG66264 standard; peptide; 8 AA.
XX
AC   AAG66264;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #56.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   3
FT                   /label= Abu, Nle
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              ::|||:|:|
Db          1 HAXRRLIF 8


RESULT 25
AAG66265
ID   AAG66265 standard; peptide; 8 AA.
XX
AC   AAG66265;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #57.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;
```

```
  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HALRRLIF 8


RESULT 26
AAG65145
ID   AAG65145 standard; peptide; 8 AA.
XX
AC   AAG65145;
XX
DT   21-NOV-2001  (first entry)
XX
DE   Synthetic peptide, p21 N-terminus-LIF hybrid.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21 N-terminus-LIF hybrid.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /label= OTHER
FT                   /note= "Other= unidentified"
FT   Modified-site   8
FT                   /note= "C-terminal carboxamide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 9; Page 48; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a synthetic peptide derived
CC   from the N-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              ::||:|:|
Db          1 KAXRRLIF 8


RESULT 27
AAG66207
ID   AAG66207 standard; peptide; 8 AA.
XX
AC   AAG66207;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived peptide, p21(152)Ser153Ala #7.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21(152)Ser153Ala.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Misc-difference 7
FT                   /label= Ile, Ala, OTHER, Leu, Val, Nle, Phe, Nva, 1Nal
FT                   /note= "Other= Cyclohexylalanine or absent"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 19; Page 57; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
```

```
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21 and used in a Cyclin A binding experiment, the effect
CC   on cyclin A binding of replacing the Ile residue at position 7 was
CC   assessed
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|||
Db          1 HAKRRLXF 8


RESULT 28
AAG66263
ID   AAG66263 standard; peptide; 8 AA.
XX
AC   AAG66263;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #55.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAARRLIF 8


RESULT 29
AAG65127
ID   AAG65127 standard; peptide; 8 AA.
XX
AC   AAG65127;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21WAFI C-terminal peptide #30.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   8
FT                   /note= "Optional C-terminal carboxamide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 15; Page 84; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
```

```
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide corresponding to
CC   p21(145-164) or a peptide derived from that region
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HSKRRLIF 8


RESULT 30
AAU05706
ID   AAU05706 standard; protein; 8 AA.
XX
AC   AAU05706;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #73.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLAF 8


RESULT 31
AAU05741
ID   AAU05741 standard; protein; 8 AA.
XX
AC   AAU05741;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #110.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 34; Page 91; 102pp; English.
XX
```

```
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 FAKRRLIF 8


RESULT 32
AAG66253
ID   AAG66253 standard; peptide; 8 AA.
XX
AC   AAG66253;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #45.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other= 3-Pyridylalanine, 2-thienylalanine or
FT                   diaminobutyric acid. The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              |::|:|:|
Db          1 XAKRRLIF 8


RESULT 33
AAG66272
ID   AAG66272 standard; peptide; 8 AA.
XX
AC   AAG66272;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #64.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
```

```
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRNLIF 8


RESULT 34
AAG66203
ID   AAG66203 standard; peptide; 8 AA.
XX
AC   AAG66203;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived peptide, p21(152)Ser153Ala #3.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21(152)Ser153Ala.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Misc-difference 3
FT                   /label= lys, Ala, Nle, Abu, Leu, Arg
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 15; Page 55; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21 and used in a Cyclin A binding experiment, the effect
CC   on cyclin A binding of replacing the Lys residue at position 3 was
CC   assessed
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              ::||:|:|
Db          1 HAXRRLIF 8


RESULT 35
AAU05740
ID   AAU05740 standard; protein; 8 AA.
XX
AC   AAU05740;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #109.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "Other= Thiophenylalanine, diaminobutyric acid or
FT                   homoserine"
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
```

```
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 34; Page 91; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              |::|:|:|
Db          1 XAKRALIF 8


RESULT 36
AAG66205
ID   AAG66205 standard; peptide; 8 AA.
XX
AC   AAG66205;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived peptide, p21(152)Ser153Ala #5.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   p21(152)Ser153Ala.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Misc-difference 5
FT                   /label= Arg, Ala, OTHER, Asn, Pro, Ser, Aib
FT                   /note= "Other= Citrulline or sarcosine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 17; Page 56; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21 and used in a Cyclin A binding experiment, the effect
CC   on cyclin A binding of replacing the Arg residue at position 5 was
CC   assessed
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|||:|
Db          1 HAKRXLIF 8


RESULT 37
AAU05709
ID   AAU05709 standard; protein; 8 AA.
XX
AC   AAU05709;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #76.
```

```
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   7
FT                   /label= OTHER, Nle, Nva, 1Nal
FT                   /note= "Other= cyclohexlalanine"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 88; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   50.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    4;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|||
Db          1 HAKRRLXF 8


RESULT 38
AAG65150
ID   AAG65150 standard; peptide; 8 AA.
XX
AC   AAG65150;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 derived cyclin A binding peptide #2.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening;
KW   mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Optionally a D-form residue"
FT   Modified-site   1
FT                   /note= "Hydrogenated N-terminus"
FT   Misc-difference 2
FT                   /note= "Optionally a D-form residue"
FT   Misc-difference 3
FT                   /note= "Optionally a D-form residue"
FT   Misc-difference 4
FT                   /note= "Optionally a D-form residue"
FT   Misc-difference 5
FT                   /note= "Optionally a D-form residue"
FT   Misc-difference 7
FT                   /note= "Optionally a D-form residue"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
FT   Misc-difference 8
FT                   /note= "Optionally a D-form residue"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Example 12; Page 53; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
```

```
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-termius of p21 and used in a Cyclin A binding experiment, the effect on
CC   cyclin A binding of replacing each residue with its chiral alternative
CC   was tested
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HAKRRLIF 8


RESULT 39
AAG66266
ID   AAG66266 standard; peptide; 8 AA.
XX
AC   AAG66266;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #58.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /note= "C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 87; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HARRRLIF 8


RESULT 40
AAJ02281
ID   AAJ02281 standard; peptide; 8 AA.
XX
AC   AAJ02281;
XX
DT   02-JUL-2001  (first entry)
XX
DE   Hepatitis C virus epitope #2272.
XX
KW   Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW   antiviral.
XX
OS   Hepatitis C virus.
XX
PN   WO200121189-A1.
XX
PD   29-MAR-2001.
XX
PF   19-JUL-2000; 2000WO-US019774.
XX
PR   19-JUL-1999;   99US-00357737.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S,  Livingston BD,  Chesnut R;
PI   Baker DM,  Celis E,  Kubo RT,  Grey HM;
XX
DR   WPI; 2001-308046/32.
XX
PT   A new composition useful as a vaccines against hepatitis C virus.
XX
PS   Disclosure; Page 157; 214pp; English.
XX
CC   The present invention describes a composition comprising a prepared
CC   hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC   These are derived from HCV HLA-binding motifs. They are useful in
CC   vaccines for the prevention and treatment of HCV infection in humans. The
CC   present sequence is an epitope used in the disclosure of the invention
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 KGGRHLIF 8


RESULT 41
ADJ72146
```

```
ID   ADJ72146 standard; peptide; 8 AA.
XX
AC   ADJ72146;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Cyclin recognition motif (CRM) from human p21 (C) protein.
XX
KW   Poly-adenylate polymerase; poly-A polymerase; PAP;
KW   cyclin recognition motif; CRM; dividing cell; tumour; non-dividing cell;
KW   cell death; cytostatic; human; p21.
XX
OS   Homo sapiens.
XX
PN   US6696546-B1.
XX
PD   24-FEB-2004.
XX
PF   06-NOV-2000; 2000US-00707263.
XX
PR   06-NOV-2000; 2000US-00707263.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Bond GL,  Manley JL,  Prives C;
XX
DR   WPI; 2004-212375/20.
XX
PT   New peptides spanning poly-adenylate polymerase's cyclin recognition
PT   motif, useful for killing dividing cells, treating abnormalities and
PT   tumors in a subject, and protecting non-dividing cells from cell death.
XX
PS   Disclosure; SEQ ID NO 9; 27pp; English.
XX
CC   The present invention relates to peptides derived from poly-adenylate
CC   (poly-A) polymerase (PAP) cyclin recognition motif (CRM). Also disclosed
CC   is a pharmaceutical composition comprising the peptides cited above and a
CC   pharmaceutical carrier, and methods of killing dividing cells and
CC   treating abnormalities and tumours in a subject using the peptides. The
CC   peptides are useful for killing dividing cells, treating abnormalities
CC   and tumours in a subject, and protecting non-dividing cells from cell
CC   death. The peptides are also useful for preparing a pharmaceutical
CC   composition for treating an abnormality. The present sequence represents
CC   a CRM from a human protein.
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 8;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 HSKRRLIF 8


RESULT 42
ADJ72150
ID   ADJ72150 standard; peptide; 8 AA.
XX
AC   ADJ72150;
XX
DT   06-MAY-2004  (first entry)
XX
DE   Cyclin recognition motif (CRM) from human Cdc25A protein.
XX
KW   Poly-adenylate polymerase; poly-A polymerase; PAP;
KW   cyclin recognition motif; CRM; dividing cell; tumour; non-dividing cell;
KW   cell death; cytostatic; human; Cdc25A.
XX
OS   Homo sapiens.
XX
PN   US6696546-B1.
XX
PD   24-FEB-2004.
XX
PF   06-NOV-2000; 2000US-00707263.
XX
PR   06-NOV-2000; 2000US-00707263.
XX
PA   (UYCO ) UNIV COLUMBIA NEW YORK.
XX
PI   Bond GL,  Manley JL,  Prives C;
XX
DR   WPI; 2004-212375/20.
XX
PT   New peptides spanning poly-adenylate polymerase's cyclin recognition
PT   motif, useful for killing dividing cells, treating abnormalities and
PT   tumors in a subject, and protecting non-dividing cells from cell death.
XX
PS   Disclosure; SEQ ID NO 13; 27pp; English.
XX
CC   The present invention relates to peptides derived from poly-adenylate
CC   (poly-A) polymerase (PAP) cyclin recognition motif (CRM). Also disclosed
CC   is a pharmaceutical composition comprising the peptides cited above and a
CC   pharmaceutical carrier, and methods of killing dividing cells and
CC   treating abnormalities and tumours in a subject using the peptides. The
CC   peptides are useful for killing dividing cells, treating abnormalities
CC   and tumours in a subject, and protecting non-dividing cells from cell
CC   death. The peptides are also useful for preparing a pharmaceutical
CC   composition for treating an abnormality. The present sequence represents
CC   a CRM from a human protein.
XX
SQ   Sequence 8 AA;

  Query Match             100.0%;  Score 20;  DB 8;  Length 8;
  Best Local Similarity   37.5%;  Pred. No. 2e+06;
  Matches    3;  Conservative    5;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 XXXRXLXF 8
              :::|:|:|
Db          1 PAPRRLLF 8


RESULT 43
ADZ77243
ID   ADZ77243 standard; peptide; 8 AA.
XX
AC   ADZ77243;
XX
DT   14-JUL-2005  (first entry)
XX
DE   Rapamycin derivative #3.
XX
KW   cytostatic; antipsoriatic; vasotropic; immunostimulant;
KW   immunosuppressive; protein interaction; cell proliferation; cell cycle;
KW   hyperproliferation; cancer; cytostatic; neoplasm; growth disorder;
KW   hyperplasia; psoriasis; antipsoriatic; dermatological disease;
KW   immune disorder; vasotropic; restenosis; cardiovascular disease;
KW   immune disorder; autoimmune disease; graft versus host disease;
KW   immunosuppressive; rapamycin.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= OTHER
FT                   /note= "OTHER= NH2-N-Im(trityl)"
FT   Modified-site   8
FT                   /label= OTHER
FT                   /note= "OTHER= CH2OH, D form residue"
XX
PN   WO2005042567-A1.
XX
PD   12-MAY-2005.
XX
```

PF 03-NOV-2004; 2004WO-CA001918.
XX
PR 03-NOV-2003; 2003US-0516273P.
XX
PA (ALTA-) ALTACHEM PHARMA LTD.
XX
PI Sharma SK, Woo T, Naicker S;
XX
DR WPI; 2005-356159/36.
XX
PT New rapamycin peptide conjugate compounds are H3 thymidine uptake
PT inhibitors useful for the treatment of e.g. cancer, hyperplasia,
PT psoriasis and hyperproliferative vascular disease.
XX
PS Example 18; Page 37; 60pp; English.
XX
CC The invention describes rapamycin peptide conjugate compounds (I). Also
CC described are: a stent coated with (I); and a composition comprising (I).
CC (I) are useful for the treatment of cell proliferation disorders e.g.
CC cancer, hyperplasia, psoriasis and hyperproliferative vascular disease
CC (restenosis), immunological condition, autoimmune disease and host-graft
CC disease. (I) is useful as immunosuppressant. This is the amino acid
CC sequence of a rapamycin derivative created in the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRRLIF 8

RESULT 44
ADZ71454
ID ADZ71454 standard; peptide; 8 AA.
XX
AC ADZ71454;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #39.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 11; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRRLIF 8

RESULT 45
ADZ71458
ID ADZ71458 standard; peptide; 8 AA.
XX
AC ADZ71458;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #43.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.

XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 12; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 KACRRLIF 8

RESULT 46
ADZ71490
ID ADZ71490 standard; peptide; 8 AA.
XX
AC ADZ71490;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #75.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 12; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a

CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRSLIF 8

RESULT 47
ADZ71491
ID ADZ71491 standard; peptide; 8 AA.
XX
AC ADZ71491;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #76.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 12; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 50.0%; Pred. No. 2e+06;
Matches 4; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|||:|
Db 1 HAKRXLIF 8

RESULT 48
ADZ71494
ID ADZ71494 standard; peptide; 8 AA.
XX
AC ADZ71494;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #79.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX

PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 12; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRRLIF 8

RESULT 49
ADZ71504
ID ADZ71504 standard; peptide; 8 AA.
XX
AC ADZ71504;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #89.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 13; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;

CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRRLAF 8

RESULT 50
ADZ71838
ID ADZ71838 standard; peptide; 8 AA.
XX
AC ADZ71838;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #423.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 9; Page 56; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 100.0%; Score 20; DB 9; Length 8;
Best Local Similarity 37.5%; Pred. No. 2e+06;
Matches 3; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Qy 1 XXXRXLXF 8
:::|:|:|
Db 1 HAKRRLIF 8

Search completed: May 5, 2006, 12:22:09
Job time : 189 secs